

(19) World Intellectual Property Organization   
International Bureau

(43) International Publication Date  
27 December 2002 (27.12.2002)

(10) International Publication Number  
**WO 02/102820 A1**

(51) International Patent Classification<sup>1</sup>: C07H 19/06. (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, IE, ES, FI, GB, GI, GH, GM, IR, IU, ID, IL, IN, IS, JP, KE, KG, KR, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PT, RO, RU, SD, SE, SG, SI, SK, SL, TL, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(21) International Application Number: PCT/DK02/00420  
(22) International Filing Date: 20 June 2002 (2006/2002)

(25) Filing Language: English  
(26) Publication Language: English

(30) Priority Data: PA 2001 00962 20 June 2001 (2006/2001) DK  
(71) Applicant (for all designated States except US): NUEVOLITION A/S (DK/DK); Rønnebægade 8, 5th floor, DK-2100 Copenhagen Ø (DK).

(72) Inventor; and  
(73) Inventor/Applicants (for US only): PEDERSEN, Henrik (DK/DK); Fredersvej 24, DK-2830 Bagsværd (DK). — with international search report before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments  
BILGAARD SLØK, Frank (DK/DK); Jagtvej 15, 3.rv., DK-2200 København N (DK). GODSESEN, Michael, Anders (DK/DK); Plantagervej 8, DK-2950 Vedbæk (DK). HYLDTØFT, Lene (DK/DK); Solskærmarken 21, DK-2830 Virum (DK). KJARNER, Sams, Christian (DK/DK); Jacob Dammefeldsvej 4A, 1., DK-1973 Frederiksberg C (DK).

(74) Title: NUCLEOSIDE DERIVATIVES FOR LIBRARY PREPARATION

(52) Description: *For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

15

The present invention relates to nucleotide derivatives. The nucleotide derivatives of the present invention are useful in the preparation of templated molecules.

5

Recently, a number of procedures have been suggested that should allow a more

10

efficient generation and screening of a larger number of molecules. The approaches

15

taken involve the encoding and/or templating of molecules other than natural bio-

polymers such as peptide, RNA and DNA. These approaches allow the researcher

20

to generate and screen a huge number of molecules in a short time. This should

25

lead to better molecules carrying the desired properties.

30

The central dogma of biology describes the one-way flow of information from DNA to

35

RNA to protein. Recently, methods such as phage display, peptides-on-plasmids,

30

ribosome display and mRNA-protein fusion have been developed, allowing the

35

transfer of information from the level of protein/peptide to RNA or DNA. This has

30

enabled the use of molecular evolution to be applied on huge numbers of peptides

35

that are exposed to an enrichment process, where after the enriched pool of mole-

30

cules (enriched for a particular feature, such as binding to receptor protein) are am-

35

plified, by exploiting information flow from the peptide to DNA and then amplifying

the DNA.

More recently, approaches have been developed that allow the encoding of poly-

30

peptides and other biochemical polymers. An example of this approach is disclosed

35

in US 5,723,598, which pertains to the identification of a biochemical polymer that

30

participates in a preselected binding interaction with a target to form a binding reac-

35

tion complex. The prior art method encompasses the generation of a library of bi-

30

functional molecules. One part of the bifunctional molecule is the biochemical poly-

35

mer and the other part is an identifier oligonucleotide comprising a sequence of nu-

30

cleotides which encodes and identifies the biochemical polymer. Following the gen-

35

eration of the library of the bifunctional molecules, a partitioning with respect to affin-

ity towards the target is conducted and the identifier oligonucleotide part of the bi-

30

ity towards the target is conducted and the identifier oligonucleotide part of the bi-

35

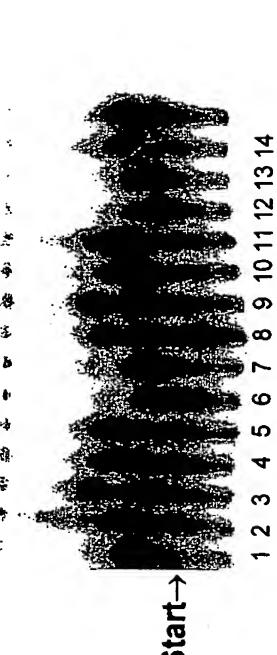
ity towards the target is conducted and the identifier oligonucleotide part of the bi-

30

ity towards the target is conducted and the identifier oligonucleotide part of the bi-

35

(57) Abstract: Nucleoside derivatives as building blocks for templated libraries are described.



**WO 02/102820 A1**

functional molecule is amplified by means of PCR. Eventually, the PCR amplicons are sequenced and decoded for identification of the biochemical polymer. This approach does not, however, allow one-pot amplification of the library members. Thus the flow of information from the identifier sequence to the biochemical polymer is restrained.

Halpin and Harbury have in WO 00/23458 suggested an improvement to the approach stipulated immediately above, wherein the molecules formed are not only identified but also directed by the nucleic acid tag. The approach is based on the traditional split-and-combine strategy for synthesis of combinatorial libraries comprising two or more synthetic steps. Plurality nucleic acid templates are used, each having at one end a chemical reactive site and dispersed throughout the strand a plurality of codons regions, each of said codon regions in turn specifying different codons. Separately, each of the strands, identified by a first codon region, is reacted at the chemical reaction sites with specific selected reagents. Subsequently, all the strands are pooled and subjected to a second partitioning based on a second codon region. The split-and-combine method is conducted an appropriate number of times to produce a library of typically between  $10^3$  and  $10^6$  different compounds. The split-and-combine method is cumbersome and generates only a relatively small library.

5

Halpin and Harbury have in WO 00/23458 suggested an improvement to the approach stipulated immediately above, wherein the molecules formed are not only identified but also directed by the nucleic acid tag. The approach is based on the traditional split-and-combine strategy for synthesis of combinatorial libraries comprising two or more synthetic steps. Plurality nucleic acid templates are used, each having at one end a chemical reactive site and dispersed throughout the strand a plurality of codons regions, each of said codon regions in turn specifying different codons. Separately, each of the strands, identified by a first codon region, is reacted at the chemical reaction sites with specific selected reagents. Subsequently, all the strands are pooled and subjected to a second partitioning based on a second codon region. The split-and-combine method is conducted an appropriate number of times to produce a library of typically between  $10^3$  and  $10^6$  different compounds. The split-and-combine method is cumbersome and generates only a relatively small library.

10

Halpin and Harbury have in WO 00/23458 suggested an improvement to the approach stipulated immediately above, wherein the molecules formed are not only identified but also directed by the nucleic acid tag. The approach is based on the traditional split-and-combine strategy for synthesis of combinatorial libraries comprising two or more synthetic steps. Plurality nucleic acid templates are used, each having at one end a chemical reactive site and dispersed throughout the strand a plurality of codons regions, each of said codon regions in turn specifying different codons. Separately, each of the strands, identified by a first codon region, is reacted at the chemical reaction sites with specific selected reagents. Subsequently, all the strands are pooled and subjected to a second partitioning based on a second codon region. The split-and-combine method is conducted an appropriate number of times to produce a library of typically between  $10^3$  and  $10^6$  different compounds. The split-and-combine method is cumbersome and generates only a relatively small library.

15

Halpin and Harbury have in WO 00/23458 suggested an improvement to the approach stipulated immediately above, wherein the molecules formed are not only identified but also directed by the nucleic acid tag. The approach is based on the traditional split-and-combine strategy for synthesis of combinatorial libraries comprising two or more synthetic steps. Plurality nucleic acid templates are used, each having at one end a chemical reactive site and dispersed throughout the strand a plurality of codons regions, each of said codon regions in turn specifying different codons. Separately, each of the strands, identified by a first codon region, is reacted at the chemical reaction sites with specific selected reagents. Subsequently, all the strands are pooled and subjected to a second partitioning based on a second codon region. The split-and-combine method is conducted an appropriate number of times to produce a library of typically between  $10^3$  and  $10^6$  different compounds. The split-and-combine method is cumbersome and generates only a relatively small library.

20

Halpin and Harbury have in WO 00/23458 suggested an improvement to the approach stipulated immediately above, wherein the molecules formed are not only identified but also directed by the nucleic acid tag. The approach is based on the traditional split-and-combine strategy for synthesis of combinatorial libraries comprising two or more synthetic steps. Plurality nucleic acid templates are used, each having at one end a chemical reactive site and dispersed throughout the strand a plurality of codons regions, each of said codon regions in turn specifying different codons. Separately, each of the strands, identified by a first codon region, is reacted at the chemical reaction sites with specific selected reagents. Subsequently, all the strands are pooled and subjected to a second partitioning based on a second codon region. The split-and-combine method is conducted an appropriate number of times to produce a library of typically between  $10^3$  and  $10^6$  different compounds. The split-and-combine method is cumbersome and generates only a relatively small library.

25

The various known methods for production of libraries as well as novel not yet public methods of the present applicant require building blocks comprising a complementing element able to recognize a coding element of a template. The present invention aims at providing such building blocks. In one aspect, the present invention relates to building blocks capable of being incorporated by a polymerase or reverse transcriptase. In another aspect, the present invention relates to building blocks capable of being incorporated in the absence of an enzyme. The building block comprises, apart from the complementing element, a linker and a functional entity. The functional entity of the compounds of the present invention may comprise an amino acid precursor. When a plurality of the building blocks are incorporated into a complementing template the functional entities are able to be linked to each other, thus forming a templated molecule, the synthesis of which is directed by the coding elements of the template. The characteristic alkyne moiety of the linkers of the present invention makes it possible to display the functional entity in the major groove of a double stranded molecule. When two or more functional entities are displayed simultaneously in the major groove reactive groups of the functional entities

30

35

simultaneously in the major groove reactive groups of the functional entities may react, either directly or via a suitable bridging molecule, to form a linkage between the functional entities. Thus, upon proper incorporation of a plurality of the compounds of the invention it is possible to form a templated molecule by linking each of the functional entities. The linkers may optionally be cleaved simultaneously with or after the formation of the templated molecule. Preferably at least one linker remains uncleaved to attach the templated molecule to the template which templated the synthesis thereof or a complementing template. A library of different complexes of template (or complementing template) and templated molecule may be subjected to various screening methods, such as affinity screening, known to the person skilled in the art to identify one or more templated molecule with the desired effect.

The compounds of the present invention may be used for the production of natural  $\alpha$ -peptides. However, recently a strong interest has been observed in academic societies for peptides other than  $\alpha$ -peptides, such as  $\beta$ -peptides,  $\gamma$ -peptides, and  $\delta$ -peptides. In one aspect of the invention it is contemplated to provide building blocks for the formation of molecules based on such artificial peptides.

#### Summary of the Invention

20 The present invention relates to nucleoside derivatives of the general formula:



Wherein Y is a group  $-\text{X}-\text{R}^2-\text{C}\equiv\text{C}-\text{N}\text{S}$   
 Wherein:  
 X is a hetero atom selected from the group O, S, Se or a group NR', wherein R' is hydrogen or an optionally substituted linear or branched C<sub>1-6</sub> alkyl or C<sub>2-6</sub> alkenyl.  
 25 R' is selected from the group consisting of C<sub>1-6</sub> alkyl, C<sub>2-6</sub> alkyliden, C<sub>2-6</sub> alkynyl, C<sub>3-6</sub> cycloalkyl, heterocycloalkyl, -CH<sub>2</sub>-O, aryl or heteroaryl, wherein each of the groups R' are substituted with 0-3 R<sup>b</sup> groups independently selected from =O, =S, -F, -Cl, -Br, -I, -OCH<sub>3</sub>, -NO<sub>2</sub> or C<sub>1-6</sub> alkyl, and  
 30 Ns is a nucleoside analogue consisting of a nucleobase and a backbone unit,

4

or Y is  $-OR^3$ , wherein R<sup>3</sup> is H or an acid protective group

R(S) is a C<sub>1-4</sub> alkylen, C<sub>3-10</sub> cycloalkylen, aryl, heterocycloalkylen or heteroaryl substituted by n sidechains S, wherein n is an integer of 0 to 4

R<sup>1</sup> is H, C<sub>1-4</sub> alkyl substituted with 0-3 R<sup>9</sup> where R<sup>9</sup> is independently selected from =O, Cl, Br, -CN, -OR<sup>6</sup>, -SR<sup>7</sup>, -NR<sup>8</sup>R<sup>7</sup>, -COOR<sup>8</sup>, -CONR<sup>9</sup>R<sup>7</sup>, -SO<sub>2</sub>NR<sup>6</sup>R<sup>7</sup> or a C<sub>1-4</sub> alkyl group forming a ringstructure with S

R<sup>6</sup> and R<sup>7</sup> are independently selected from H, C<sub>1-4</sub> linear alkyl, C<sub>1-4</sub> branched alkyl, C<sub>1-4</sub> cycloalkyl, aryl, heteroaryl, aralkyl, or hetero aralkyl.

S is C<sub>1-4</sub> linear alkyl, C<sub>3-4</sub> branched alkyl, C<sub>3-6</sub> cycloalkyl, aryl, heteroaryl, aralkyl, hetero aralkyl substituted with 0-3 R<sup>5</sup> where R<sup>5</sup> is independently selected from =O, Cl, Br, -CN, -OR<sup>6</sup>, -SR<sup>7</sup>, -NR<sup>8</sup>R<sup>7</sup>, -COOR<sup>8</sup>, -CONR<sup>9</sup>R<sup>7</sup>, -SO<sub>2</sub>NR<sup>6</sup>R<sup>7</sup>.

Z is H, an amino protective group or a group  $\text{---C---R}^2\text{---C}\equiv\text{C---Ns}$  with the proviso, that when Y is not  $\text{---X---R}^2\text{---C}\equiv\text{C---Ns}$ , Z is  $\text{---C---R}^2\text{---C}\equiv\text{C---Ns}$

Such derivatives enable the preparation of large libraries of compounds templated by nucleic acids or analogues thereof. In particular, the present invention relates to building blocks carrying amino acid components allowing the construction of oligopeptides containing natural- as well as unnatural amino acid fragments.

In a preferred embodiment the alkynylene linker is connected to the nucleobase of a nucleoside analogue.

In another preferred embodiment the alkynylene linker is connected to the nucleobase of a nucleoside analogue in the 7 position of the bicyclic purine nucleobases and the 5 position of the monocyclic pyrimidine bases which ensures the positioning of the functional entity into the major groove of the nascent oligomer-complex.

The combination of R<sup>2</sup> and X determines the stability of the linkage between the functional entity and the complementing element. Hence different R<sup>2</sup>-X combinations require different cleavage conditions allowing some linkers to be cleaved while others remain intact.

5

In a preferred embodiment R<sup>2</sup> is selected from the group consisting of C<sub>1-4</sub> alkyl, C<sub>2-6</sub> alkylarylen, C<sub>2-6</sub> alkynylene, heterocycloalkylen, -CH<sub>2</sub>O-, aryl or heteroaryl, each of the groups R<sup>2</sup> are substituted with 0-3 R<sup>8</sup> groups independently selected from =O, -F, -Cl, -Br, -NO<sub>2</sub>, C<sub>1-4</sub> alkyl.

In a preferred embodiment R<sup>2</sup> is selected from the group consisting of C<sub>1-4</sub> alkyl, C<sub>2-6</sub> alkylarylen, heterocycloalkylen, -CH<sub>2</sub>O-, aryl or heteroaryl/en, each of the groups R<sup>2</sup> are substituted with 0-2 R<sup>8</sup> groups independently selected from =O, -F, -NO<sub>2</sub>, C<sub>1-4</sub> alkyl.

In a preferred embodiment R<sup>2</sup> is selected from the group consisting of -CH<sub>2</sub>-, -CH<sub>2</sub>CH<sub>2</sub>-, -CH<sub>2</sub>O-, or aryl/en each of the groups R<sup>2</sup> are substituted with 0-2 R<sup>8</sup> groups independently selected from =O, -F, -NO<sub>2</sub>, C<sub>1-4</sub> alkyl.

In a preferred embodiment R<sup>2</sup> is selected from the group consisting of -CH<sub>2</sub>-, -CH<sub>2</sub>CH<sub>2</sub>-, -CH<sub>2</sub>O-, or aryl/en each of the groups R<sup>2</sup> are substituted with 0-2 R<sup>8</sup> groups independently selected from =O, -F, -NO<sub>2</sub>, C<sub>1-4</sub> alkyl.

In a preferred embodiment R<sup>2</sup> is selected from the group consisting of -CH<sub>2</sub>-, -CH<sub>2</sub>CH<sub>2</sub>-, -CH<sub>2</sub>O-, or aryl/en.

In a preferred embodiment R<sup>2</sup> is selected from the group consisting of -CH<sub>2</sub>-, -CH<sub>2</sub>CH<sub>2</sub>-, -CH<sub>2</sub>O-, or aryl/en.

In a preferred embodiment R<sup>2</sup> is selected from the group consisting of -CH<sub>2</sub>-, -CH<sub>2</sub>CH<sub>2</sub>-, -CH<sub>2</sub>O-, or aryl/en.

In a preferred embodiment R<sup>2</sup> is selected from the group consisting of -CH<sub>2</sub>-, -CH<sub>2</sub>CH<sub>2</sub>-, -CH<sub>2</sub>O-, or aryl/en.

In a preferred embodiment R<sup>2</sup> is selected from the group consisting of -CH<sub>2</sub>-, -CH<sub>2</sub>CH<sub>2</sub>-, -CH<sub>2</sub>O-, or aryl/en.

In a preferred embodiment R<sup>2</sup> is selected from the group consisting of -CH<sub>2</sub>-, -CH<sub>2</sub>CH<sub>2</sub>-, -CH<sub>2</sub>O-, or aryl/en.

In a preferred embodiment R<sup>2</sup> is selected from the group consisting of -CH<sub>2</sub>-, -CH<sub>2</sub>CH<sub>2</sub>-, -CH<sub>2</sub>O-, or aryl/en.

In a preferred embodiment R<sup>2</sup> is selected from the group consisting of -CH<sub>2</sub>-, -CH<sub>2</sub>CH<sub>2</sub>-, -CH<sub>2</sub>O-, or aryl/en.

In a preferred embodiment R<sup>2</sup> is selected from the group consisting of -CH<sub>2</sub>-, -CH<sub>2</sub>CH<sub>2</sub>-, -CH<sub>2</sub>O-, or aryl/en.

In a preferred embodiment R<sup>2</sup> is selected from the group consisting of -CH<sub>2</sub>-, -CH<sub>2</sub>CH<sub>2</sub>-, -CH<sub>2</sub>O-, or aryl/en.

In a preferred embodiment R<sup>2</sup> is selected from the group consisting of -CH<sub>2</sub>-, -CH<sub>2</sub>CH<sub>2</sub>-, -CH<sub>2</sub>O-, or aryl/en.

In a preferred embodiment R<sup>2</sup> is selected from the group consisting of -CH<sub>2</sub>-, -CH<sub>2</sub>CH<sub>2</sub>-, -CH<sub>2</sub>O-, or aryl/en.

In a preferred embodiment R<sup>2</sup> is selected from the group consisting of -CH<sub>2</sub>-, -CH<sub>2</sub>CH<sub>2</sub>-, -CH<sub>2</sub>O-, or aryl/en.

In a preferred embodiment R<sup>2</sup> is selected from the group consisting of -CH<sub>2</sub>-, -CH<sub>2</sub>CH<sub>2</sub>-, -CH<sub>2</sub>O-, or aryl/en.

In a preferred embodiment R<sup>2</sup> is selected from the group consisting of -CH<sub>2</sub>-, -CH<sub>2</sub>CH<sub>2</sub>-, -CH<sub>2</sub>O-, or aryl/en.

In a preferred embodiment R(S) is a C<sub>1-4</sub> alkylene, C<sub>3-10</sub> cycloalkylene, aryl or heteroaryl or heteroaryl substituted by n sidechains S, wherein n is an integer of 0 to 3

In a preferred embodiment R(S) is a C<sub>1-4</sub> alkylene, aryl or heteroaryl substituted by n sidechains S, wherein n is an integer of 0 to 3

6

In a preferred embodiment R(S) is a C<sub>1-4</sub> alkylene substituted by n sidechains S, wherein n is an integer of 0 to 3

In a preferred embodiment R(S) is a C<sub>1-2</sub> alkylene substituted by n sidechains S, wherein n is an integer of 0 to 3

5 In a preferred embodiment R(S) is a C<sub>1-2</sub> alkylene substituted by n sidechains S, wherein n is an integer of 0 to 2

In a preferred embodiment R(S) is a C<sub>1-2</sub> alkylene substituted by n sidechains S, wherein n is an integer of 0 to 1

In a preferred embodiment S is C<sub>1-4</sub> linear alkyl, C<sub>3-4</sub> branched alkyl, C<sub>2-6</sub> cycloalkyl, ary, heteroaryl, aralkyl, hetero aralkyl substituted with 0-3 R<sup>6</sup> where R<sup>6</sup> is independently selected from =O, Cl, Br, -CN, -OR<sup>6</sup>, -SR<sup>6</sup>, -NR<sup>6</sup>R<sup>7</sup>, -CONRR<sup>7</sup>, -SO<sub>2</sub>NR<sup>6</sup>R<sup>7</sup> where R<sup>6</sup> and R<sup>7</sup> are independently selected from H, C<sub>1-3</sub> linear alkyl, C<sub>2-6</sub> cycloalkyl, ary, heteroaryl, aralkyl, or hetero aralkyl.

In a preferred embodiment S is C<sub>1-4</sub> linear alkyl, C<sub>3-4</sub> branched alkyl, C<sub>2-6</sub> cycloalkyl, ary, heteroaryl, aralkyl, hetero aralkyl substituted with 0-2 R<sup>6</sup> where R<sup>6</sup> is independently selected from =O, Cl, -CN, -OR<sup>6</sup>, -SR<sup>6</sup>, -NR<sup>6</sup>R<sup>7</sup>, -CONRR<sup>7</sup>, -SO<sub>2</sub>NR<sup>6</sup>R<sup>7</sup> where R<sup>6</sup> and R<sup>7</sup> are independently selected from H, C<sub>1-3</sub> linear alkyl, ary, heteroaryl, aralkyl, or hetero aralkyl.

In a preferred embodiment S is C<sub>1-4</sub> linear alkyl, C<sub>3-4</sub> branched alkyl, C<sub>2-6</sub> cycloalkyl, ary, heteroaryl, aralkyl, hetero aralkyl substituted with 0-2 R<sup>6</sup> where R<sup>6</sup> is independently selected from =O, Cl, -CN, -OR<sup>6</sup>, -SR<sup>6</sup>, -NR<sup>6</sup>R<sup>7</sup>, -CONRR<sup>7</sup>, -SO<sub>2</sub>NR<sup>6</sup>R<sup>7</sup> where R<sup>6</sup> and R<sup>7</sup> are independently selected from H and C<sub>1-3</sub> linear alkyl.

In a preferred embodiment S is C<sub>1-4</sub> linear alkyl, C<sub>3-4</sub> branched alkyl, C<sub>2-6</sub> cycloalkyl, ary, heteroaryl, aralkyl, hetero aralkyl substituted with 0-1 R<sup>6</sup> where R<sup>6</sup> is selected from =O, Cl, -CN, -OR<sup>6</sup>, -SR<sup>6</sup>, -NR<sup>6</sup>R<sup>7</sup>, -CONRR<sup>7</sup>, -SO<sub>2</sub>NR<sup>6</sup>R<sup>7</sup> where R<sup>6</sup> and R<sup>7</sup> are independently selected from H and C<sub>1-3</sub> linear alkyl.

In a preferred embodiment S is C<sub>1-4</sub> linear alkyl or aryl substituted with 0-1 R<sup>6</sup> where R<sup>6</sup> is selected from =O, Cl, -CN, -OR<sup>6</sup>, -SR<sup>6</sup>, -NR<sup>6</sup>R<sup>7</sup>, -CONRR<sup>7</sup>, -SO<sub>2</sub>NR<sup>6</sup>R<sup>7</sup> where R<sup>6</sup> and R<sup>7</sup> are independently selected from H and C<sub>1-3</sub> linear alkyl.

SUBSTITUTE SHEET (RULE 26)

WO 02/102820 7

In a preferred embodiment R<sup>1</sup> is H, C<sub>1-6</sub> alkyl substituted with 0-1 R<sup>6</sup> where R<sup>6</sup> is independently selected from =O, Cl, Br, -CN, -OR<sup>6</sup>, -SR<sup>6</sup>, -NR<sup>6</sup>R<sup>7</sup>, -COOR<sup>6</sup>, -CONRR<sup>6</sup>R<sup>7</sup>, -SO<sub>2</sub>NR<sup>6</sup>R<sup>7</sup> R<sup>6</sup> and R<sup>7</sup> are independently selected from H, C<sub>1-6</sub> linear alkyl, C<sub>1-6</sub> branched alkyl, C<sub>1-6</sub> cycloalkyl, ary, heteroaryl, aralkyl, or hetero aralkyl, or a C<sub>1-6</sub> alkylene substituted by n sidechains S.

5 In a preferred embodiment R<sup>1</sup> is H, C<sub>1-6</sub> alkyl or a C<sub>1-6</sub> alkylene substituted by n sidechains S.

In a preferred embodiment R<sup>1</sup> is H or a C<sub>1-6</sub> alkylene substituted by n sidechains S.

In a preferred embodiment R<sup>1</sup> is H or a C<sub>1-6</sub> alkylene substituted by n sidechains S.

In a preferred embodiment R<sup>1</sup> is H or a C<sub>1-6</sub> alkylene substituted by n sidechains S.

10 In a preferred embodiment R<sup>1</sup> is H or a C<sub>1-6</sub> alkylene substituted by n sidechains S.

In a preferred embodiment R<sup>1</sup> is H.

In a preferred embodiment Z is H, an amino protective group selected from the group of formyl, acetyl, trifluoroacetyl, benzoyl, trifluoroxyacetyl, triphenylmethyl, benzyl, benzyloxyacetyl or tosyl or a group  $\text{---C}(\text{---R}^2\text{---C}\equiv\text{C---Ns})$  with the proviso, that when Y is not  $\text{---X---R}^2\text{---C}\equiv\text{C---Ns}$ , then Z is  $\text{---C}(\text{---R}^2\text{---C}\equiv\text{C---Ns})$ .

15 In a preferred embodiment Z is H, an amino protective group selected from the group of acetyl, trifluoroacetyl, benzyl, benzyloxyacetyl, tosyl or a group  $\text{---C}(\text{---R}^2\text{---C}\equiv\text{C---Ns})$  with the proviso, that when Y is not  $\text{---X---R}^2\text{---C}\equiv\text{C---Ns}$ , then Z is  $\text{---C}(\text{---R}^2\text{---C}\equiv\text{C---Ns})$ .

20 In a preferred embodiment the nucleobase is uracil or cytosine modified in the 5 position or 7-deazaguanidine or 7-deazaadenine or 2'-deazaguanidine modified in the 7 position.

25 In a preferred embodiment the backbone unit type is DNA, RNA, Oxy-LNA, Thio-LNA, Amino-LNA, Phosphorthioate, 2'-O-methyl, PNA or Morpholino as described in chart 3.

30 In a preferred embodiment the backbone unit type is DNA, RNA, Oxy-LNA, PNA or Morpholino.

SUBSTITUTE SHEET (RULE 26)

In a preferred embodiment the backbone unit type is DNA, PNA or Oxy-LNA

In a preferred embodiment the backbone unit type is DNA

In a preferred embodiment the backbone unit type is Oxy-LNA

In a preferred embodiment the backbone unit type is PNA

5 Using di- or trimeric building blocks results in improved recognition of the nucleobases on the template, especially when chemical methods are used to oligomerise the nucleoside analogues. (Schmidt; 1997, *Nucleic Acids Research*; 4792-4796) The use of oligomeric nucleoside analogues allow the direct annealing of building blocks to the template without the need for chemical- or enzymatic incorporation.

In a preferred embodiment more nucleoside analogues are connected via their backbone structures forming di-, tri- or oligomeric nucleoside analogues as building blocks

10 In a preferred embodiment Y is —X—R<sup>2</sup>—C≡C—N<sup>3</sup>S or -OR<sup>3</sup>, wherein R<sup>3</sup> is selected from the group H, C<sub>1-3</sub> alkyl, allyl, benzyl, *tert*-butyl or triphenylmethyl.

15 Aralkyl is an aryl connected to a C<sub>1-3</sub> alkylene Complementing element recognizes combinations of nucleobases in the template and consists of at least one nucleoside analogue, optionally attached to a series of at least one backbone unit carrying a nucleobase.

Complex is a templated molecule linked to the template that templated the synthesis of the templated molecule. The template can be a complementing template as defined herein that is optionally hybridised or otherwise attached to a corresponding template of linked coding elements.

20 Heteraryl designates an unsaturated cyclic structure consisting of 2-5 carbon atoms and 1-3 heteroatoms selected from O, S, N or P.

Heterocycloalkyl designates a saturated or partially saturated cyclic structure consisting of 2-5 carbon atoms and 1-3 heteroatoms selected from O, S, N or P.

Library is in this context a collection of molecules.

Nucleoside analogue is any combination of a nucleobase and a backbone unit.

	Abbreviations
5	DCC DIC DEA DMAP EDC HATU
10	HOAt HOBI NHS PyBOP
15	TBTU TEA

#### Brief description of the charts

In chemical structure drawings throughout the document, hydrogen atoms on terminal carbon atoms are not explicitly shown.

#### Detailed Description of the Invention

Building blocks consist apart from a linker and a functional entity of one or more nucleoside analogues i.e. pairs of nucleobases and backbone units, forming the complementing entity and may as such be considered a nucleoside derivative.

The nucleobase may be of natural or of synthetic origin but all shares the common feature of being able to selectively recognize one other nucleobase. Examples of such base pairs are shown in chart 1

10

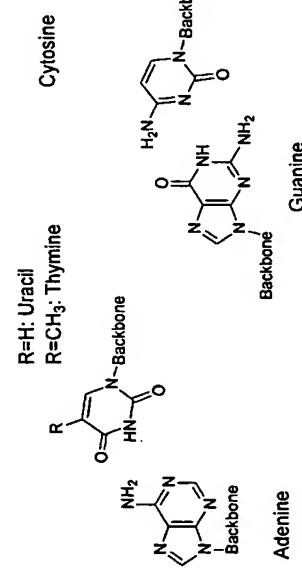
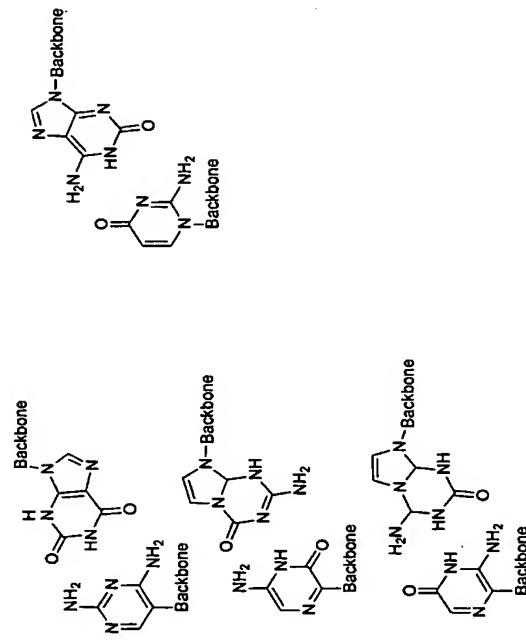
**Natural Base Pairs****Synthetic Base Pairs**

Chart 1 Natural and Synthetic nucleobases.

Also, modifications to both natural- and synthetic nucleobases is possible without obliteration of the mutual recognition properties, e.g. replacing the N7 atom of adenine or guanine with a carbon atom affords 7-deaza adenine and 7-deaza guanine

5

respectively (Chart 2) that still recognises natural thymine or uracil and cytosine, respectively. Further the introduction of substituents at certain positions on the complementing entity is also possible.

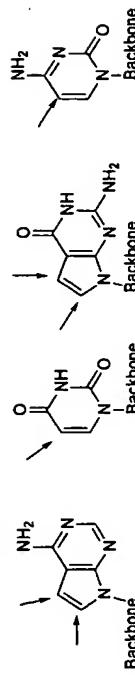
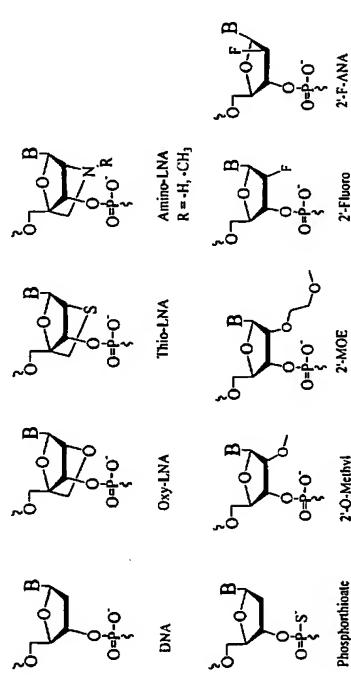


Chart 2, 7-deaza-adenine, uracil, 7-deaza-guanidine and cytosine. Arrows indicate preferred sites of substitution on the nucleobase of the complementing entity that direct the functional entity into the major groove of the nascent oligomer complex.

The backbone units of the building blocks may contain a set of reactive groups that enables enzymatic or chemical oligomerisation of the building blocks. Examples of backbone unit structures are given in chart 3

10

12



13

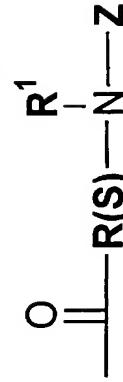
The linker is based on a rigid alkynylene spacer that positions the functional entity

away from the back bone of the oligomer complex:  $—X—R^2—C\equiv C—$

X is a hetero atom selected from the group O, S, Se or a group NR<sup>4</sup>, wherein R<sup>4</sup> is hydrogen or an optionally substituted linear or branched C<sub>1-4</sub> alkyl or C<sub>2-8</sub> alkenyl.

5 R<sup>2</sup> is selected from the group consisting of C<sub>1-4</sub> alkylen, C<sub>2-4</sub> alkylenealkylen, C<sub>2-4</sub> alkylen, C<sub>3-4</sub> cycloalkylen, heterocycloalkylen, -CH<sub>2</sub>-O-, arylene or heteroarylen, wherein each of the groups R<sup>2</sup> are substituted with 0-3 R<sup>8</sup> groups independently selected from =O, =S, =F, -Cl, -Br, -I, -OCH<sub>3</sub>, -NO<sub>2</sub> or C<sub>1-4</sub> alkyl

10 The functional entity is an aminoacid derivative:



Wherein:

R(S) is a C<sub>1-4</sub> alkylen, C<sub>3-10</sub> cycloalkylen, aryl, heterocycloalkylen or heteroaryl substituted by n sidechains S, wherein n is an integer of 0 to 4

5 R<sup>1</sup> is H, C<sub>1-4</sub> alkyl substituted with 0-3 R<sup>5</sup> where R<sup>5</sup> is independently selected from =O, Cl, Br, -CN, -OR<sup>6</sup>, -SR<sup>6</sup>, -COOR<sup>6</sup>, -CONR<sup>6</sup>R<sup>7</sup>, -SO<sub>2</sub>NR<sup>6</sup>R<sup>7</sup> or a C<sub>1-4</sub> alkylen group forming a ringsstructure with S

R<sup>6</sup> and R<sup>7</sup> are independently selected from H, C<sub>1-4</sub> linear alkyl, C<sub>1-4</sub> branched alkyl, C<sub>1-4</sub> cycloalkyl, aryl, heteroaryl, aralkyl, or hetero aralkyl.

15 S is C<sub>1-4</sub> linear alkyl, C<sub>2-8</sub> branched alkyl, C<sub>3-4</sub> cycloalkyl, aryl, heteroaryl, aralkyl, hetero aralkyl substituted with 0-3 R<sup>5</sup> where R<sup>5</sup> is independently selected from =O, Cl, Br, -CN, -OR<sup>6</sup>, -SR<sup>6</sup>, -NR<sup>6</sup>R<sup>7</sup>, -COOR<sup>6</sup>, -CONR<sup>6</sup>R<sup>7</sup>, -SO<sub>2</sub>NR<sup>6</sup>R<sup>7</sup>.

Z is H, an amino protective group

#### General Synthesis Procedures

The compounds of the invention are generally prepared by two different methods.

5 Building blocks may be oligomerised using enzymatic or chemical methods. (Schmidt; 1997, *Nucleic Acids Research*; 4792-4796; Inoue; 1984, *Journal of Molecular Biology*; 669-676; Schmidt; 1997, *Nucleic Acids Research*; 4797-4802) Enzymatic incorporation is typically based on the use of 5'-O-triphosphate building blocks with a ribose derived backbone unit. Chemical incorporation of building blocks with a ribose derived backbone unit relies on the use of an activated phosphate ester e.g. a phosphorimidate. (Zhao, 1998, *J. Org. Chem.*; 7568-7572) For peptide backbone units, peptide coupling reagents are employed.

As shown in chart 3 several modifications of the natural DNA- and RNA backbone is

possible, particularly the 2'-position of the ribose entity is well suited for functional

entity linkage.

15

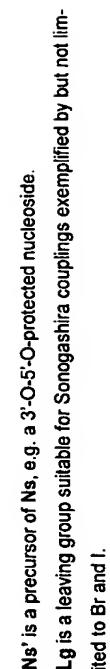
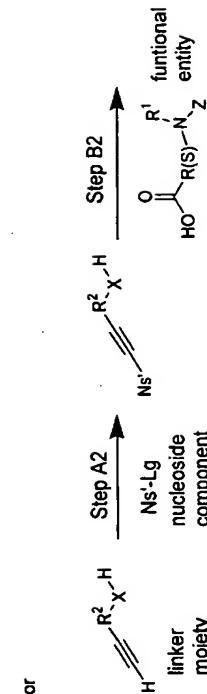
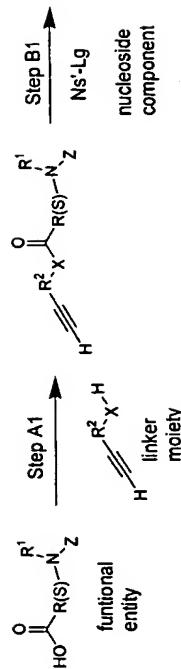
Chart 3 Backbone units used and building blocks. B designates the nucleobase and wavy bonds show points of oligomerisation.

5 Building blocks may be oligomerised using enzymatic or chemical methods. (Schmidt; 1997, *Nucleic Acids Research*; 4792-4796; Inoue; 1984, *Journal of Molecular Biology*; 669-676; Schmidt; 1997, *Nucleic Acids Research*; 4797-4802) Enzymatic incorporation is typically based on the use of 5'-O-triphosphate building blocks with a ribose derived backbone unit. Chemical incorporation of building blocks with a ribose derived backbone unit relies on the use of an activated phosphate ester e.g. a phosphorimidate. (Zhao, 1998, *J. Org. Chem.*; 7568-7572) For peptide backbone units, peptide coupling reagents are employed.

As shown in chart 3 several modifications of the natural DNA- and RNA backbone is

possible, particularly the 2'-position of the ribose entity is well suited for functional

14



Ns' is a precursor of Ns, e.g. a 3'-O-5'-O-protected nucleoside.

Lg is a leaving group suitable for Sonogashira couplings exemplified by but not limited to Br and I.

Step A1

The amino acid derivative (functional entity) (10.37 mmol) is dissolved in a solvent exemplified by but not limited to dichloromethane, 1,2-dichloroethane, 1,2-dichloropropane, tetrahydrofuran, dimethylformamide or a mixture hereof and added a peptide coupling reagent (12.44 mmol, 1.2 eq) exemplified by but not limited to EDC, DCC, DIC, HATU, PyBOP or PyBOP optionally in the presence of a peptide coupling enhancer like HOBr, HOAt, or NHS at a temperature of -20-100 °C preferably 0-50 °C. To this mixture, the linker moiety (15.55 mmol, 1.5 equiv) is added optionally in the presence of DMAP (1.04 mmol, 0.1 eq) and the reaction is

Step B1

left 2-16 h. Upon evaporation of volatiles, the residue is taken up in dichloromethan and washed with HCl (aq, 0.1 M); NaHCO<sub>3</sub> (aq, sat); and water. Removal of dichloromethan affords the crude product which is further purified by chromatography if necessary.

5

Step B2

A solution of the nucleoside component (0.34 mmol) in a solvent like dimethylformamide, dimethylsulfoxid, toluene, tetrahydrofuran, water, ethanol, methanol or a mixture hereof is added a terminal alkyne (the linker moiety-functional entity) (0.69 mmol, 2 eq) and a base like DIEA (0.25 mL) and is purged with Ar for 5 min. Tetakis triphenylphosphine palladium (0.03 mmol, 0.1 eq) and CuI (0.07 mmol, 0.2 eq) is added and the reaction is run at 20-100 °C, preferably at 20-50 °C, and kept there for 20 h. Evaporation of volatiles followed by chromatography affords the desired modified nucleoside.

10

Step A2

A solution of the complementing element precursor (0.34 mmol) in a solvent like dimethylformamide, dimethylsulfoxid, toluene, tetrahydrofuran, water, ethanol, methanol or a mixture hereof is added a terminal alkyne (the linker moiety) (0.69 mmol, 2 eq) and a base like DIEA (0.25 mL) and is purged with Ar for 5 min. Tetakis triphenylphosphine palladium (0.03 mmol, 0.1 eq) and CuI (0.07 mmol, 0.2 eq) is added and the reaction is run at 20-100 °C, preferably at 20-50 °C, and kept there for 20 h. Evaporation of volatiles followed by chromatography affords the desired modified nucleoside.

15

Step B2

Depending on the nature of Ns' several steps known from literature may be required to convert Ns' into Ns e.g. Protective group removal (Greene, 1999; ) or conversion of 5'OH groups of nucleosides into 5'0-riphosphates or phosphorimidazolides.(Zhao, 1998; J. Org. Chem.; 75:58-7572)

20

Nucleoside analogues with phosphate linkages in the backbone may be combined with wild type nucleotides to form di-, tri- or oligomeric buildingblocks. Likewise, nucleoside analogues having a PNA backbone unit may be combined with PNA monomers to form di-, tri- or oligomeric building blocks.

30

16

The amino acid derivative (functional entity) (10.37 mmol) is dissolved in a solvent exemplified by but not limited to dichloromethane, 1,2-dichloroethane, 1,2-dichloropropane, tetrahydrofuran, dimethylformamid or a mixture hereof and added a peptide coupling reagent (12.44 mmol, 1.2 eq) exemplified by but not limited to EDC, DCC, DIC, HATU, HBTU, PyBOP or PyBOP optionally in the presence of a peptide coupling enhancer like HOBr, HOAt, or NHS at a temperature of -20-100 °C preferably 0-50 °C. To this mixture, the linker-nucleoside component (15.55 mmol, 1.5 equiv) obtained in step A2 is added optionally in the presence of DMAP (1.04 mmol, 0.1 eq) and the reaction is left 2-16 h. Upon evaporation of volatiles, the residue is taken up in dichloromethan and washed with HCl (aq, 0.1 M); NaHCO<sub>3</sub> (aq, sat); and water. Removal of dichloromethan affords the crude product which may be further purified by chromatography if necessary.

Depending on the nature of N's' several steps known from literature may be required to convert N's into N's, e.g. protective group removal, conversion of 5'-OH groups of ribose derived backbone units into 5'-O-triphosphates or phosphorimidazolides. (Zhao: 1998; *J. Org. Chem.*; 7568-7572). For peptide derived backbone units other types of modifications are required. (Hyup: 1996; *Bioorganic & medicinal chemistry*, 5-23)

20

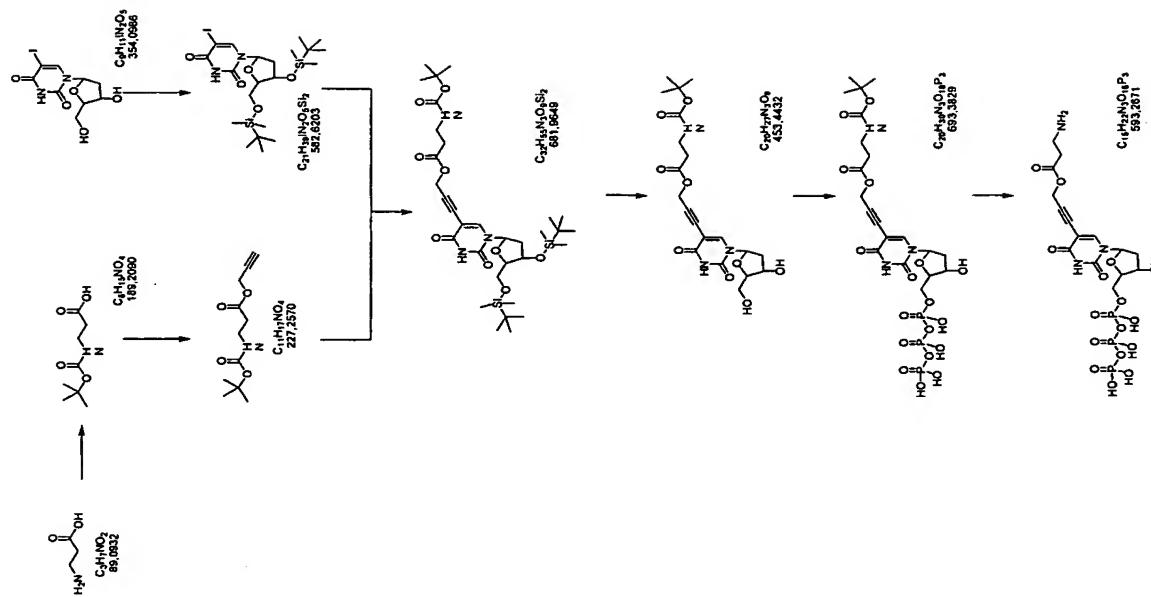
Nucleoside analogues carrying a ribose derived backbone unit may be combined with wild type nucleotides to form di-, tri- or oligo-nucleotid building blocks. Likewise, nucleoside analogues having a peptide backbone unit may be combined with PNA monomers to form di-, tri or oligo peptidic building blocks.

25

### Examples

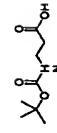
#### Example 1 to 7: Preparation of the mononucleotide building block (I)

Building block I may be prepared according to the general scheme shown below:



17

**Example 1: Preparation of 3-*tert*-Butyloxycarbonylamino-propionic acid (*N*-Boc- $\beta$ -alanine)(1a)**



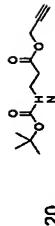
5

To a solution of  $\beta$ -alanine (2.25 g, 25 mmol) in aq.  $\text{NaHCO}_3$  (25 mL) were added *tert*-butyl dicarbonate (4.35 g, 20 mmol) and acetonitrile (25 mL). The reaction mixture was stirred at room temperature for 18 h.

10  $\text{EtOAc}$  (100 mL) was added and pH was adjusted to 4.5 by addition of  $\text{NaH}_2\text{PO}_4$ . The product was extracted into  $\text{EtOAc}$  (3  $\times$  50 mL), dried ( $\text{Na}_2\text{SO}_4$ ), and evaporated to dryness under vacuum to afford 3.71 g (98%).

15  $^1\text{H NMR}$  ( $\text{CDCl}_3$ )  $\delta$  11 (1H, br s, COOH), 5.07 (1H, br s, NH), 3.40 (2H, m), 2.58 (2H, m), 1.44 (9H, s,  $^3\text{Bu}$ ).

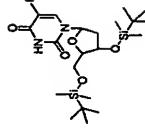
**Example 2: Preparation of *N*-Boc- $\beta$ -alanine propargyl ester(1b).**



20

*N*-Boc- $\beta$ -alanine (1.91 g, 10.1 mmol) and propargyl alcohol (0.675 g, 12 mmol) were dissolved in  $\text{EtOAc}$  (25 mL). Dicyclohexyl-carbodiimide (DCC, 2.06 g, 10 mmol) was added to the solution and after 16 h of stirring at room temperature, the reaction mixture was filtered and evaporated to dryness under vacuum. Crude product yield

**Example 3: Preparation of 5-*iodo*-2'-*deoxyuridine* 3',5'-Di-*tert*-butyldimethylsilyl Ether(1c).**

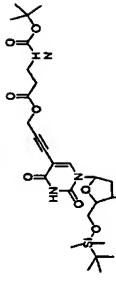


5 5-*iodo*-2'-*deoxyuridine* (Aldrich, 2.39 g, 6.7 mmol) and imidazole (2.025 g, 29.7 mmol) was dissolved in anhydrous  $\text{DMF}$  (10 mL). A solution of *tert*-butyldimethylsilyl chloride (2.24 g, 14.9 mmol) in anhydrous  $\text{DMF}$  (5 mL) was added and the resulting mixture was stirred for 16 h at room temperature.

10 The reaction mixture was poured into  $\text{EtOAc}$  (400 mL), washed with  $\text{NH}_4\text{Cl}$  (50% sat. aq, 80 mL) followed by water (80 mL). After drying with  $\text{Na}_2\text{SO}_4$ ,  $\text{EtOAc}$  was removed under reduced pressure to leave a colorless oil that solidified on standing. Recrystallization in n-hexane (14 mL) afforded 2.64 g, 80%.

15  $^1\text{H NMR}$  ( $\text{CDCl}_3$ )  $\delta$  8.18 (1H, br s, NH); 8.10 (1H, s); 6.23 (1H, dd); 4.40 (1H, dd); 4.05 (1H, dd); 3.92 (1H, dd); 3.78 (1H, dd); 2.32 (1H, dd); 2.05 (1H, ddd); 0.95 (9H, s,  $^3\text{Bu}$ ); 0.90 (9H, s,  $^3\text{Bu}$ ); 0.15 (3H, s,  $\text{CH}_3$ ); 0.13 (3H, s,  $\text{CH}_3$ ); 0.08 (3H, s,  $\text{CH}_3$ ); 0.07 (3H, s,  $\text{CH}_3$ ).

**Example 4: Preparation of compound (1d)**



Compound (1d)

20

20

A solution of iodo silyl ether (1c) (1.62 g, 2.7 mmol), *N*-Boc- $\beta$ -alanine (1a) (2.03 g, 8.9 mmol) and triethylamine (0.565 g, 5.8 mmol) in 10 mL dry DMF were stirred at room temperature.  $\text{N}_2$  was passed through the solution for 20 min.

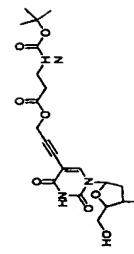
5 Tetrakis(triphenylphosphine)palladium(0) (269 mg, 0.2 mmol) and copper(I) iodide (90 mg, 0.4 mmol) were added and the reaction mixture was stirred at room temperature for 32 h.

EtOAc (100 mL) was poured into the reaction mixture, followed by washing (aq  $\text{NaHCO}_3$  (50 mL); brine (50 mL)), drying ( $\text{Na}_2\text{SO}_4$ ), and removal of solvent by vacuum evaporation.

10 The crude product (2.4 g) was purified by silica column chromatography eluting with EtOAc:Heptane gradient: (1:2)-(5:3) (v/v). Product yield 1.15 g, 60%.

<sup>1</sup>H NMR ( $\text{CDCl}_3$ )  $\delta$  8.45 (1H, s), 8.05 (1H, s,  $\delta\text{-H}$ ), 7.35 (1H, bs, NH), 6.25 (1H, dd, 1'-H), 4.82 (2H, s,  $\text{CH}_2\text{O}$ ), 4.39 (1H, m, 3'-H), 3.97 (1H, m, 4'-H), 3.80 (2H, dd, 5',5"-H), 3.40 (2H, m,  $\text{CH}_2\text{N}$ ), 2.58 (2H, t,  $\text{CH}_2$ ), 2.2 (1H, m, 2'-H), 2.0 (1H, m, 2'-H), 1.45 (9H, s,  $^3\text{Bu}$ ), 0.93 (9H, s,  $^3\text{Bu}$ ), 0.89 (9H, s,  $^3\text{Bu}$ ), 0.15 (3H, s,  $\text{CH}_3$ ), 0.13 (3H, s,  $\text{CH}_3$ ), 0.08 (3H, s,  $\text{CH}_3$ ), 0.07 (3H, s,  $\text{CH}_3$ ).

15 Example 5: Preparation of compound (1e)



Compound (1e)

20 A solution of *N*-Boc- $\beta$ -alanine silyl ether (1d) (100 mg, 0.15 mmol), glacial acetic acid (75 mg, 1.25 mmol) and tetrabutylammonium fluoride trihydrate (TBAF) (189 mg, 0.6 mmol) in 2 mL dry THF was stirred at room temperature for 3 d.

The reaction mixture was evaporated and purified by silica column chromatography eluting with dichloromethane(DCM):methanol(MeOH) gradient (95:5)-(88:12) (v/v).

25 Product yield 26 mg, 38%.

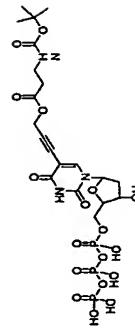
30

21

<sup>1</sup>H NMR ( $\text{CD}_3\text{OD}$ )  $\delta$  8.35 (1H, s,  $\delta\text{-H}$ ), 6.15 (1H, t, 1'-H), 4.80 (2H, s,  $\text{CH}_2\text{O}$ ), 4.32 (1H, dt, 3'-H), 3.86 (1H, q, 4'-H), 3.70 (2H, dd, 5',5"-H), 3.24 (2H, m,  $\text{CH}_2\text{N}$ ), 2.47 (2H, t,  $\text{CH}_2$ ), 2.28-2.10 (1H, m, 2',2'-H), 1.44 (9H, s,  $^3\text{Bu}$ ).

5

Example 6: Preparation of compound (1f)



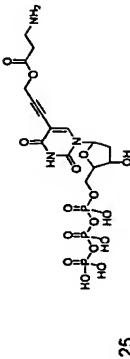
10 COMPOUND 1f

15 *N*-Boc- $\beta$ -alanine nucleoside (1e) (26 mg, 57  $\mu\text{mol}$ ) was dissolved in 200  $\mu\text{L}$  dry trimethylphosphate. After cooling to 0 °C, a solution of phosphorus oxychloride ( $\text{POCl}_3$ ) in dry trimethylphosphate was added (100  $\mu\text{L}$  stock solution (104 mg/mL), 68  $\mu\text{mol}$ ). The reaction mixture was stirred at 0 °C for 2 h.

Subsequently a solution of tributylammonium pyrophosphate (Sigma P-8533) (67.8 mg, 143  $\mu\text{mol}$  in 300  $\mu\text{L}$  dry DMF) and tributylamine (26.9 mg, 145  $\mu\text{mol}$  in 150  $\mu\text{L}$  dry DMF) was added at 0 °C. The reaction was stirred at room temperature for 3 min. and then stopped by addition of 1 mL 1.0 M triethylammonium hydrogen carbonate.

20

Example 7: Preparation of compound 1



25 COMPOUND 1

Removal of N-Boc protection group.

Following phosphorylation, 50  $\mu$ l of the phosphorylation reaction mixture is adjusted to pH = 1 using HCl and incubated at room temperature for 30 minutes. The mixture is adjusted to pH 5.5 using equimolar NaOH and Na-acetate (pH 5.5) before purification on TLC.

5

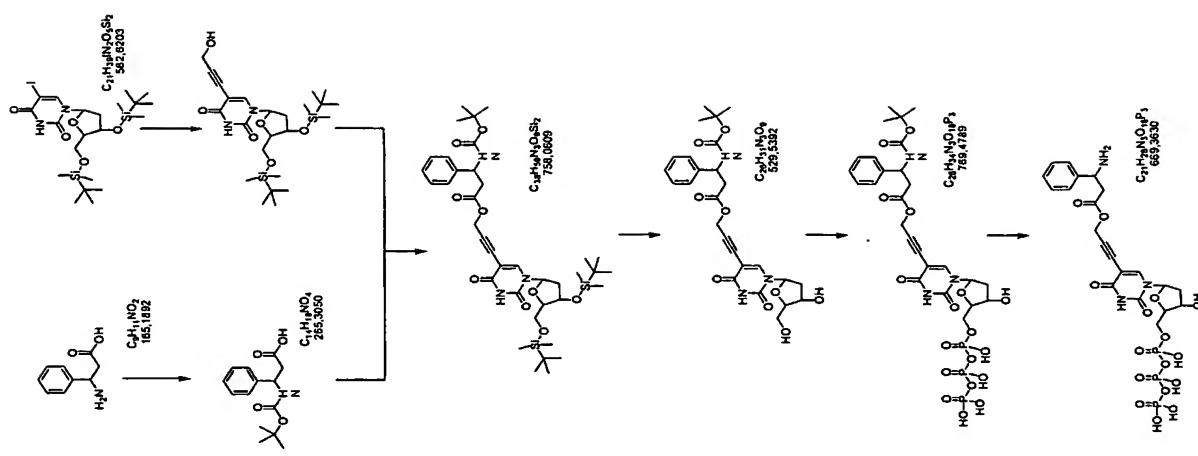
Purification of nucleotide derivatives using thin-layer chromatography (TLC)

From the crude mixture, 20 samples of 2  $\mu$ l were spotted on kieselgel 60 F<sub>254</sub> TLC (Merck). Organic solvents and non-phosphorylated nucleosides were separated from the nucleotides derivatives using 100% methanol as running solution. Subsequently, the TLC plate is air-dried and the nucleotide-derivative identified by UV-shadowing. Kiesel containing the nucleotide-derivative was isolated and extracted twice using 10 mM Na-acetate (pH = 5.5) as solvent. Kieselgel was removed by centrifugation and the supernatant was dried *in vacuo*. The nucleotide derivative was resuspended in 50-100  $\mu$ l H<sub>2</sub>O to a final concentration of 1-3 mM. The concentration of each nucleotide derivative was evaluated by UV-absorption prior to use in polymerase extension reactions.

**Examples 8 to 13: Preparation of the mononucleotide building block (II)**

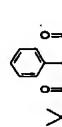
20

Building block II may be prepared according to the general scheme shown below:



### Example 8: Preparation of N-Boc-3-phenyl- $\beta$ -alanine (7a)

EtOAc (100 mL) was poured into the reaction mixture, followed by washing (aq NaHCO<sub>3</sub> (50 mL); brine (50 mL)), drying (Na<sub>2</sub>SO<sub>4</sub>), and removal of solvent by vacuum evaporation.



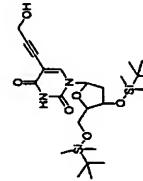
2

To a solution of 3-amino-3-phenylpropionic acid (3.30 g, 20 mmol) in  $\text{NaHCO}_3$  (50% aq. 25 mL) were added di-*tert*-butyl dicarbonate (4.36 g, 20 mmol) and acetonitrile (30 mL). The reaction mixture was stirred at room temperature for 18 h. Di-*tert*-butyl dicarbonate (4.36 g, 20 mmol) was added and the reaction mixture was stirred

at room temperature for 18 h.  $\text{NaH}_2\text{PO}_4$  was added and pH was adjusted to 4.5 by addition of  $\text{NaH}_2\text{PO}_4$ . The product was extracted into  $\text{EtOAc}$  (3  $\times$  100 mL), dried ( $\text{Na}_2\text{SO}_4$ ), and evaporated to dryness under vacuum to afford crude product 5.6 g (105%)

1

**Example 9: Preparation of 5-(3-Hydroxypropyn-1-yl)-2'-deoxyuridine 3',5'-Di-**



A solution of iodo silyl ether (3) (1.30 g, 2.2 mmol), propargyl alcohol (0.386 g, 6.9 mmol) and triethylamine (0.38 g, 4.3 mmol) in 7 ml. dry DMF was degassed with  $N_2$ . Tetakis(triphenylphosphine)palladium(0) (228 mg, 0.2 mmol) and copper(I) iodide (120 mg, 0.4 mmol) were added and the reaction mixture was stirred at room temperature for 32 h.

mmol) and 4-dimethylaminopyridine (DMAP, 10 mg) were added to the solution, and after 16 h of stirring at room temperature the reaction mixture was poured into  $\text{EtOAc}$  (100 mL), washed with  $\text{NaHCO}_3$  (50% sat. aq., 50 mL), dried ( $\text{Na}_2\text{SO}_4$ ), filtered and concentrated under vacuum.

225 The crude product was purified by silica column chromatography eluting  
EtOAc:Heptane gradient (1:2)-(2:3) (v/v). Product yield 335 mg. 88%.

<sup>1</sup>H NMR (CDCl<sub>3</sub>) δ: 8.49 (1H, s), 8.04 (1H, s, 6-H), 7.29 (5H, m, Ph), 6.27 (1H, dd, 1'-H).

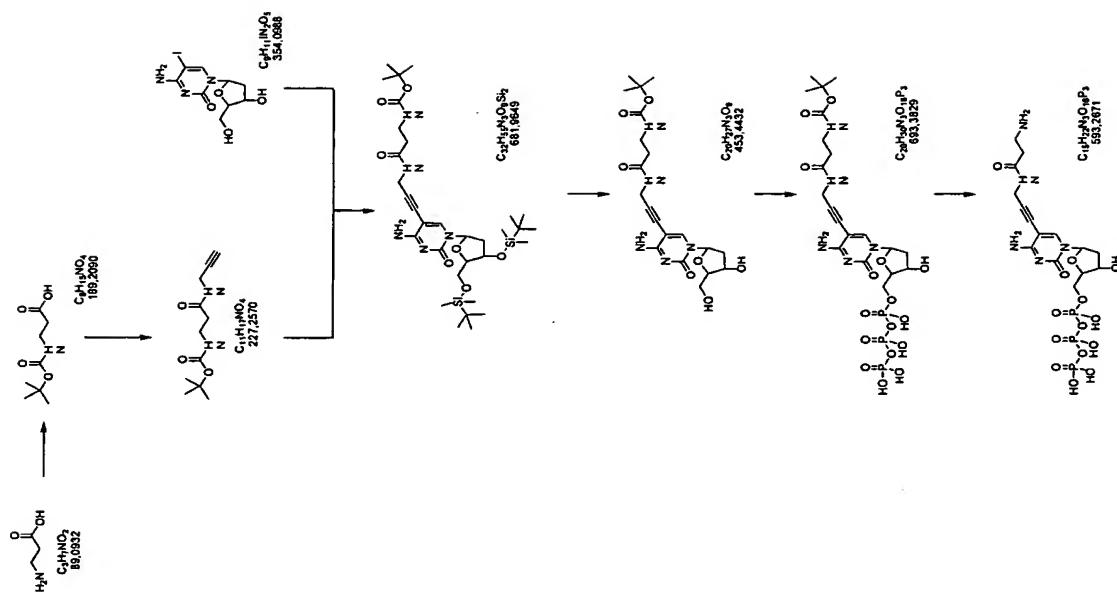


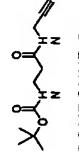
centrifugation and the supernatant was dried *in vacuo*. The nucleotide derivative was resuspended in 50–100  $\mu$ l H<sub>2</sub>O to a final concentration of 1–3 mM. The concentration of each nucleotide derivative was evaluated by UV-absorption prior to use in polymerase extension reactions.

5

**Examples 14 to 18: Preparation of the mononucleotide building block (III)**

10 Building block III may be prepared according to the general scheme shown below:



**Example 14: Preparation of *N*-Boc- $\beta$ -alanine propargyl amide (3a)**

Tetrakis(triphenylphosphine)palladium(0) (66.5 mg, 0.057 mmol) and copper(I) iodide (20.7 mg, 0.108 mmol) were added and the reaction mixture was stirred at room temperature for 5 d

Imidazole (112 mg, 1.6 mmol) was added. A solution of *tert*-butylidinemethylsilyl chloride (234 mg, 1.5 mmol) in anhydrous DMF (1 mL) was added and the resulting mixture was stirred for 16 h at room temperature.

The reaction mixture was evaporated and EtOAc (25 mL) was added. The resulting mixture was filtered and the solvent removed by vacuum evaporation.

The crude product was purified by silica column chromatography eluting with DCM:MeOH (92.5:7.5) (v/v). Product yield 84 mg, 25%.

5

10

*N*-Boc- $\beta$ -alanine (1a) (1.05 g, 5.5 mmol) and propargyl amine (0.90 g, 16.5 mmol) were dissolved in THF (10 mL). Diisopropyl-carbodiimide (DIC, 635 g, 5.5 mmol) was added and the reaction mixture was stirred for 16 h at room temperature. Water was added (20 mL) and the product was extracted into EtOAc (3x30 mL). The combined EtOAc was dried ( $\text{Na}_2\text{SO}_4$ ) and evaporated. The crude product was purified by silica column chromatography eluting with EtOAc:Heptane gradient (2:3) (3:2.5) (v/v). Product yield 0.925 g, 74 %.

10  $^1\text{H}$  NMR ( $\text{CDCl}_3$ )  $\delta$  6.69 (1H, bs, NH), 5.32 (1H, bs, NH), 4.04 (2H, bs), 3.41 (2H, dd), 2.45 (2H, t), 2.24 (1H, s), 1.44 (9H, s,  $^3\text{Bu}$ ).

20 **Example 15: Preparation of compound (3b)**

20

COMPOUND 3b

25 **Example 16: Preparation of compound (3c)**

20 **Compound 3c**

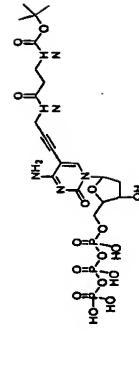
25 A solution of compound (3b) (84 mg, 0.12 mmol) and tetrabutylammonium fluoride trihydrate (TBAF) (155 mg, 0.45 mmol) in 2 mL dry THF was stirred at room temperature for 4 days.

The reaction mixture was evaporated and purified by silica column chromatography eluting with DCM:MeOH gradient (9:1)-(8:2) (v/v). Product yield 27 mg, 48%.

30  $^1\text{H}$  NMR ( $\text{CDCl}_3$ )  $\delta$  8.32 (1H, s), 6.20 (1H, dd, 1'-H), 4.35 (1H, dt), 4.15 (2H, s, CH<sub>2</sub>), 3.95 (1H, q), 3.83 (1H, dd), 3.72 (1H, dd), 3.36-3.30 (3H, m), 2.42-2.36 (3H, m), 2.13 (1H, dt), 1.40 (9H, s,  $^3\text{Bu}$ ).

is adjusted to pH 5.5 using equimolar NaOH and Na-acetate (pH 5.5) before purification on TLC.

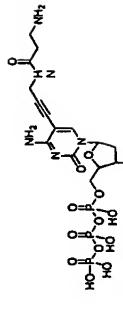
**Example 17: Preparation of compound (3d)**



**COMPOUND 3d**

Compound (3c) (27 mg, 60  $\mu$ mol) was dissolved in 100  $\mu$ L dry trimethylphosphate. After cooling to 0 °C, a solution of phosphorus oxychloride ( $\text{POCl}_3$ ) in dry trimethylphosphate was added (100  $\mu$ L stock solution (110 mg/mL), 72  $\mu$ mol). The reaction mixture was stirred at 0 °C for 2h. Subsequently a solution of triethylammonium pyrophosphate (71 mg, 150  $\mu$ mol in 300  $\mu$ L dry DMF) and triethylamine (28.3 mg, 153  $\mu$ mol in 150  $\mu$ L dry DMF) was added at 0 °C. The reaction was stirred at room temperature for 3 min. and then stopped by addition of 1.0 M triethylammonium hydrogen carbonate (1 mL).

**Example 18: Preparation of compound III**



Removal of N-Boc protection group.  
Following phosphorylation, 50  $\mu$ L of the phosphorylation reaction mixture is adjusted to pH = 1 using HCl and incubated at room temperature for 30 minutes. The mixture

**Purification of nucleotide derivatives using thin-layer chromatography (TLC)**

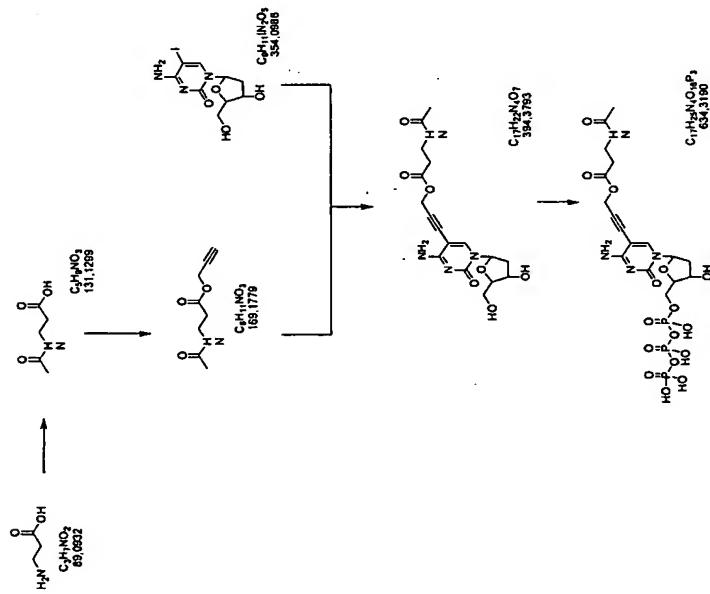
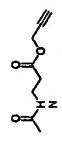
5 From the crude mixture, 20 samples of 2  $\mu$ L were spotted on Kieselgel 60 F<sub>254</sub> TLC (Merck). Organic solvents and non-phosphorylated nucleosides were separated from the nucleotides derivatives using 100% methanol as running solution. Subsequently, the TLC plate is air-dried and the nucleotide-derivative identified by UV-shadowing. Kiesel containing the nucleotide-derivative was isolated and extracted twice using 10 mM Na-acetate (pH = 5.5) as solvent. Kieselgel was removed by centrifugation and the supernatant was dried *in vacuo*. The nucleotide derivative was resuspended in 50-100  $\mu$ L H<sub>2</sub>O to a final concentration of 1-3 mM. The concentration of each nucleotide derivative was evaluated by UV-absorption prior to use in polymerase extension reactions.

10

15

20 **Examples 19 to 22: Preparation of the mononucleotide building block (IV)**

Building block IV may be prepared according to the general scheme shown below:

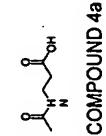
Example 20: Preparation of *N*-Acetyl- $\beta$ -alanine propargyl ester(4b).

COMPOUND 4b

To a solution of *N*-Acetyl- $\beta$ -alanine(4a) in THF (20 mL) was added propargyl alcohol (840 mg, 15 mmol), 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDC) (1.035 g, 5.39 mmol), triethylamine (540 mg, 5.4 mmol) and 4-dimethylaminopyridin (5 mg). The reaction mixture was stirred at room temperature for 2 d.

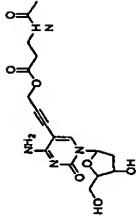
The reaction mixture was poured into EtOAc (100 mL), washed with  $\text{NaH}_2\text{PO}_4$  (50% sat. aq, 2x50 mL) followed by  $\text{NaHCO}_3$  (50% sat. aq, 50 mL). After drying ( $\text{Na}_2\text{SO}_4$ ), EtOAc was removed under reduced pressure to leave a colourless oil that solidified on standing. Product yield 536 mg, 59%.

**5** Example 19: Preparation of *N*-Acetyl- $\beta$ -alanine(4a)



COMPOUND 4a

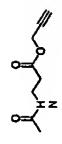
10 To a solution of  $\beta$ -alanine (2.25 g, 25 mmol) in aq.  $\text{NaHCO}_3$  (15 mL) was added acetonitrile (15 mL) and acetic anhydride (2.55 g, 25 mmol). The reaction mixture was stirred at room temperature for 3 h. Acetic anhydride (2.55 g, 25 mmol) was added and after 2 h and pH was adjusted to 4-5 by addition of  $\text{NaH}_2\text{PO}_4$ .



COMPOUND 4c

15 A solution of 5-iodo-2'-deoxyxylidin (200 mg, 0.56 mmol), triethylamine (100 mg, 1 mmol) and compound (4b) (190 mg, 1.13 mmol) in anhydrous DMF (7mL) was stirred at room temperature.  $\text{N}_2$  was passed through the solution for 20 min.

The product was extracted into EtOAc (3 x 50 mL), dried ( $\text{Na}_2\text{SO}_4$ ), and evaporated to dryness under vacuum to afford 1.96 g (60%)

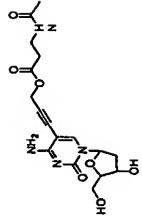
Example 20: Preparation of *N*-Acetyl- $\beta$ -alanine propargyl ester(4b).

COMPOUND 4b

To a solution of *N*-Acetyl- $\beta$ -alanine(4a) in THF (20 mL) was added propargyl alcohol (840 mg, 15 mmol), 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDC) (1.035 g, 5.39 mmol), triethylamine (540 mg, 5.4 mmol) and 4-dimethylaminopyridin (5 mg). The reaction mixture was stirred at room temperature for 2 d.

The reaction mixture was poured into EtOAc (100 mL), washed with  $\text{NaH}_2\text{PO}_4$  (50% sat. aq, 2x50 mL) followed by  $\text{NaHCO}_3$  (50% sat. aq, 50 mL). After drying ( $\text{Na}_2\text{SO}_4$ ), EtOAc was removed under reduced pressure to leave a colourless oil that solidified on standing. Product yield 536 mg, 59%.

Example 21: Preparation of compound (4c)



COMPOUND 4c

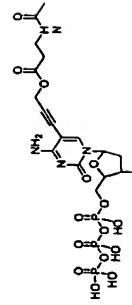
A solution of 5-iodo-2'-deoxyxylidin (200 mg, 0.56 mmol), triethylamine (100 mg, 1 mmol) and compound (4b) (190 mg, 1.13 mmol) in anhydrous DMF (7mL) was stirred at room temperature.  $\text{N}_2$  was passed through the solution for 20 min.

Tetrakis(triphenylphosphine)palladium(0) (70 mg, 0.06 mmol) and copper(I) iodide (22 mg, 0.12 mmol) were added and the reaction mixture was stirred at room temperature for 4 d.

The reaction mixture was evaporated and purified by silica column chromatography eluting with DCM:MeOH gradient (9:1-8:2) (v/v). Product yield 141 mg, 63%.

<sup>1</sup>H NMR (CD<sub>3</sub>OD) δ 8.41 (1H, s), 6.20 (1H, dd, 1'-H), 4.97 (2H, s), 4.38 (1H, d), 3.97 (1H, q), 3.85 (1H, dd), 3.75 (1H, dd), 3.46 (2H, t), 2.61 (2H, t), 2.39 (1H, m), 2.18 (1H, m).

10 Example 22: Preparation of compound IV:



15 COMPOUND IV

Compound (4c) (140 mg, 355 µmol) was dissolved in 600 µL dry trimethylphosphate. After cooling to 0 °C, a solution of phosphorus oxychloride (POCl<sub>3</sub>) in dry trimethylphosphate was added (600 µL stock solution (108 mg/mL), 420 µmol). The reaction mixture was stirred at 0 °C for 2 h.

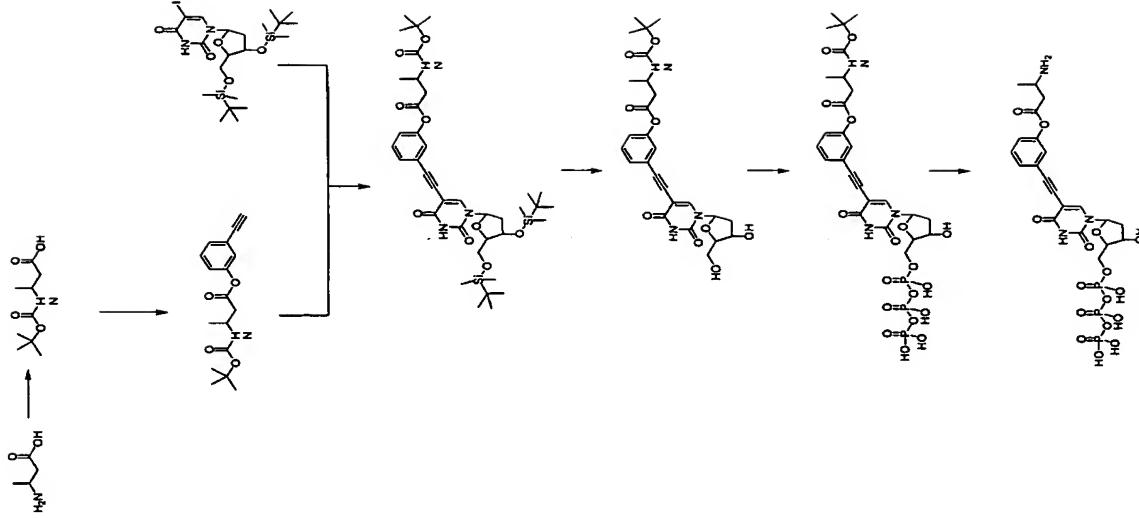
Subsequently a solution of triethylammonium pyrophosphate (422 mg, 890 µmol) in 1.8 mL dry DMF and triethylamine (168 mg, 900 µmol in 0.9 mL dry DMF) was added at 0 °C. The reaction was stirred at room temperature for 3 min. and then stopped by addition of 1.0 M triethylammonium hydrogen carbonate (1 mL).

From the crude mixture, 20 samples of 2 µL were spotted on kieselgel 60 F<sub>254</sub> TLC (Merck). Organic solvents and non-phosphorylated nucleosides were separated from the nucleotides derivatives using 100% methanol as running solution. Subsequently, the TLC plate is air-dried and the nucleotide-derivative identified by UV-shadowing. Kiesel containing the nucleotide-derivative was isolated and extracted twice using 10 mM Na-acetate (pH = 5.5) as solvent. Kieselgel was removed by

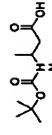
centrifugation and the supernatant was dried *in vacuo*. The nucleotide derivative was resuspended in 50-100 µL H<sub>2</sub>O to a final concentration of 1-3 mM. The concentration of each nucleotide derivative was evaluated by UV-absorption prior to use in polymerase extension reactions.

Examples 23 to 28: Preparation of the mononucleotide building block (V)

Building block V may be prepared according to the general scheme shown below:



### Example 23: Preparation of compound 5a



## COMPOUND 5a

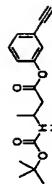
To a solution of 3-amino-butyric acid (2.06 g, 20 mmol) in  $\text{NaHCO}_3$  (50% sat. aq. 25 mL) were added di-*tert*-butyl dicarbonate (4.36 g, 20 mmol) and acetonitrile (30 mL).

10 The reaction mixture was stirred at room temperature for 18 h. Di-*t*-butyl dicar-

EtOAc (100 mL) was added and pH was adjusted to 4-5 by addition of  $\text{NaH}_2\text{PO}_4$ . The product was extracted into EtOAc (3  $\times$  100 mL), dried ( $\text{Na}_2\text{SO}_4$ ), and evapo-

The product was extracted into EtOAc ( $3 \times 100 \text{ mL}$ ),  $\text{H}_2\text{O}$  ( $100 \text{ mL}$ ), and  $\text{Et}_2\text{O}$  ( $100 \text{ mL}$ ). The EtOAc layer was collected and dried over  $\text{Na}_2\text{SO}_4$ . The product was isolated as a white solid (4.6 g, 113%).

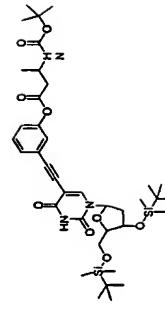
### Example 24: Preparation of compound 5b



COMPOUND 55

Compound 28 (1.023 g, 5.0 mmol), 3-Ethynyl-phenole (Lancaster, 0.675 g, 12 mmol) and 4-dimethylamino-pyridin (DMAP, 300 mg, 2.5 mmol) were dissolved in EtOAc (10 mL). Dicyclohexyl-carbodiimide (DCC, 2.06 g, 10 mmol) was added to the solution and after 16 h of stirring at room temperature, the reaction mixture was filtered and evaporated to dryness under vacuum. The crude product was purified by silica column chromatography, elution with EtOAc:Hexane gradient (1:3-1:12) (v/v).

Product yield 7220 mg, 73%.  
<sup>1</sup>H NMR (CDCl<sub>3</sub>) δ 7.36-7.09 (4H, m, Ph), 4.89 (1H, bs, NH), 4.22 (1H, br, CH), 3.10 (1H, s) 2.77 (2H, d), 1.40 (3H, t), 1.32 (3H, d)

**Example 25: Preparation of compound 5c****COMPOUND 5c**

10 A solution of 5-Iodo-2'-deoxyuridine 3',5'-Di-tert-butylidemethylsilyl ether (730 mg, 1.25 mmol), triethylamine (250 mg, 2.5 mmol) and compound(5b) (456 mg, 1.5 mmol) in anhydrous DMF (3 mL) was stirred at room temperature. N<sub>2</sub> was passed through the solution for 20 min.

15 Tetakis(triphenylphosphine)palladium(0) (109 mg, 0.094 mmol) and copper(I) iodide (36 mg, 0.188 mmol) were added and the reaction mixture was stirred at room temperature for 3 d.

The reaction mixture was evaporated and purified by silica column chromatography eluting with EtOAc:Heptane gradient (1:3)-(1:2)(v/v). Product yield 807 mg, 85%.

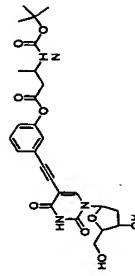
<sup>1</sup>H NMR (CDCl<sub>3</sub>) δ 8.38 (1H, s), 8.08 (1H, s, 6-H), 7.39-7.1 (4H, m, Ph), 6.33 (1H, dd, 1'-H), 4.9 (1H, bs), 4.45 (1H, dt), 4.80 (2H, s, CH<sub>2</sub>), 4.2 (1H, m), 4.02 (1H, m, 4'-H), 3.95 (1H, dd, 5'-H), 3.79 (1H, dd, 5'-H), 2.78 (2H, d), 2.36 (1H, m, 2'-H), 2.07 (1H, m, 2''-H), 1.46 (9H, s, <sup>3</sup>Bu), 0.93 (9H, s, <sup>3</sup>Bu), 0.91 (9H, s, <sup>3</sup>Bu), 0.15 (3H, s, CH<sub>3</sub>), 0.13 (3H, s, CH<sub>3</sub>), 0.11 (3H, s, CH<sub>3</sub>), 0.09 (3H, s, CH<sub>3</sub>).

20

**Example 26: Preparation of compound 5d**

25

Compound (5d) (138.5 mg, 260  $\mu$ mol) was dissolved in 500  $\mu$ L dry trimethylphosphine. After cooling to 0 °C, a solution of phosphorus oxychloride (POCl<sub>3</sub>) in dry trimethylphosphine was added (400  $\mu$ L stock solution (120 mg/mL), 310  $\mu$ mol). The reaction mixture was stirred at 0 °C for 2h.

**COMPOUND 5d**

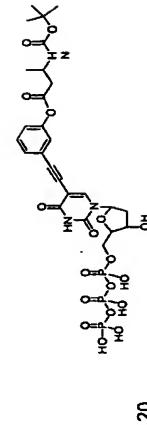
5 A solution of compound (5c) (807 mg, 1.06 mmol), glacial acetic acid (1.0 g, 16 mmol) and tetrabutylammonium fluoride trihydrate (TBAF) (2.36 g, 7.5 mmol) in 20 mL dry THF was stirred at room temperature for 3 d.

The reaction mixture was evaporated and purified by silica column chromatography eluting with (DCM):(MeOH) (9:1) (v/v). Product yield 408 mg, 72%.

10

10 <sup>1</sup>H NMR (CD<sub>3</sub>OD) δ 8.46 (1H, s, 6-H), 7.39 (2H, m, Ph), 7.28 (1H, m, Ph), 7.12 (1H, m, Ph), 6.75 (1H, bd), 6.27 (1H, dd, 1'-H), 4.44 (1H, dt, 4'-H), 3.96 (1H, t, 3'-H), 3.86 (1H, dd, 5'-H), 3.77 (1H, dd, 5'-H), 2.72 (2H, d), 2.35-2.27 (2H, m, 2', 2''-H), 1.46 (9H, s, <sup>3</sup>Bu), 1.27 (3H, d).

15

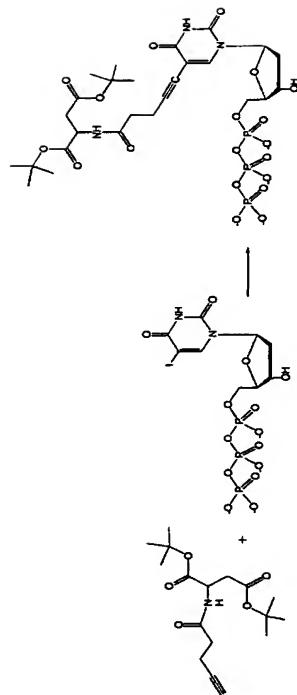
**Example 27: Preparation of compound 5e****COMPOUND 5e**

25 Compound (5d) (138.5 mg, 260  $\mu$ mol) was dissolved in 500  $\mu$ L dry trimethylphosphine. After cooling to 0 °C, a solution of phosphorus oxychloride (POCl<sub>3</sub>) in dry trimethylphosphine was added (400  $\mu$ L stock solution (120 mg/mL), 310  $\mu$ mol). The reaction mixture was stirred at 0 °C for 2h.



was pure enough for further synthesis.  $^1\text{H-NMR}$  ( $\text{CDCl}_3$ ):  $\delta$  6.6 (1H, NH), 4.6 (1H, CH), 2.8 (2H,  $\text{CH}_2$ ), 2.4 (4H,  $2 \times \text{CH}_2$ ), 1.9 (1H, CH), 1.2 (18H,  $6 \times \text{CH}_3$ ).

**Example 31: Preparation of 2-[5-[1-(4-Hydroxy-5-(O-triphosphate-hydroxymethyl)-tetrahydrofuran-2-yl)-2,4-dioxo-1,2,3,4-tetrahydro-pyrimidin-5-yl]-pent-4-ynoylamino]-succinic acid di-*tert*-butyl ester (VI)**

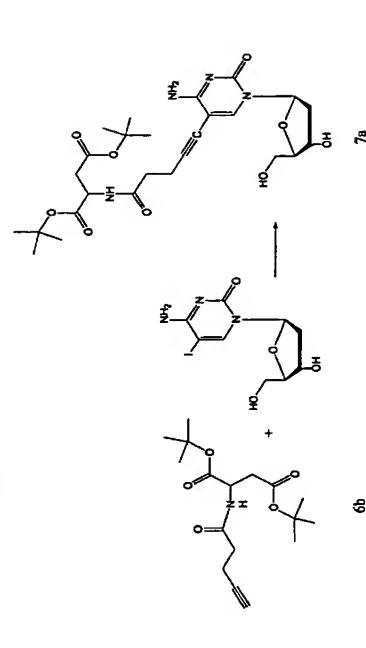


10 The nucleotide 9d (20 mg, 0.022 mmol) was dissolved in water-ethanol (1:1, 2 mL). The solution was degassed and kept under an atmosphere of argon. The catalyst  $\text{Pd}(\text{PPh}_3)_4(\text{m-C}_6\text{H}_4\text{SO}_3\text{Na})_4$  (20 mg, 0.016 mmol) prepared in accordance with A.L. Casalnuovo et al. J. Am. Chem. Soc. 1990, 112, 4324-4330, triethylamine (0.02 mL, 0.1 mmol) and the alkyne 6b (20 mg, 0.061 mmol) were added. Few crystals of Cul were added. The reaction mixture was stirred for 6 h. The triethylammonium salt of compound VI was achieved after purification by RP-HPLC (eluent: 100mM triethylammonium acetate  $\rightarrow$  20% acetonitrile in 100mM triethylammonium acetate),  $^1\text{H-NMR}$  ( $\text{D}_2\text{O}$ ):  $\delta$  8.1 (1H, CH), 6.2 (1H, CH), 4.8 (1H, CH), 4.6 (1H, CH), 4.1 (3H,  $\text{CH}_2$ ), 2.8 (2H,  $\text{CH}_2$ ), 2.7 (2H,  $\text{CH}_2$ ), 2.5 (2H,  $\text{CH}_2$ ), 2.3 (2H,  $\text{CH}_2$ ), 1.4 (18H,  $6 \times \text{CH}_3$ ).

15 Immediately prior to incorporation or after incorporation, the protective di-*tert*-butyl ester groups may be cleaved to obtain the corresponding free carboxylic acid.

**Examples 32 to 33: Preparation of the mononucleotide building block (VII)**

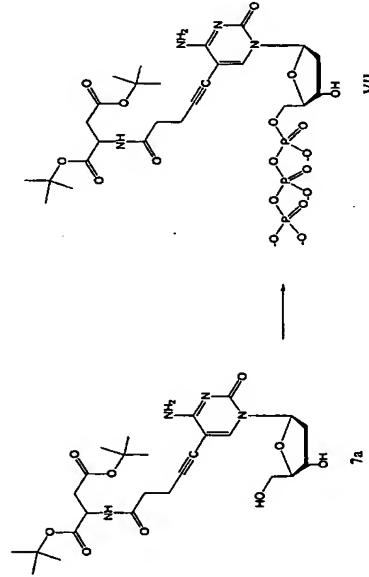
**Example 32: Preparation of 2-[5-[4-Amino-1-(4-hydroxy-5-hydroxymethyl-tetrahydrofuran-2-yl)-2-oxo-1,2-dihydro-pyrimidin-5-yl]-pent-4-ynoylamino]-succinic acid di-*tert*-butyl ester (7a)**



10 Compound (7a) (30 mg, 19%) was obtained from compound (6b) (140 mg, 0.43 mmol) and 5-iodo-2-deoxycytidine (100 mg, 0.28 mmol) using the procedure described for the synthesis of compound VI.  $^1\text{H-NMR}$  (MeOD-D<sub>4</sub>):  $\delta$  8.3 (1H, CH), 6.2 (1H, CH), 4.8 (1H, CH), 4.6 (1H, CH), 4.4 (1H, CH), 4.0 (1H, CH), 3.8 (2H,  $\text{CH}_2$ ), 2.8 (4H, 2  $\times$   $\text{CH}_2$ ), 2.7 (2H,  $\text{CH}_2$ ), 2.4 (1H,  $\text{CH}_2$ ), 2.2 (1H,  $\text{CH}_2$ ), 1.4 (18H, 6  $\times$   $\text{CH}_3$ ).

15 **Example 32: Preparation of 2-[5-[4-Amino-1-(4-hydroxy-5-(O-triphosphate-hydroxymethyl)-tetrahydro-furan-2-yl)-2-oxo-1,2-dihydro-pyrimidin-5-yl]-pent-4-ynoylamino]-succinic acid di-*tert*-butyl ester (Compound VII)**

20



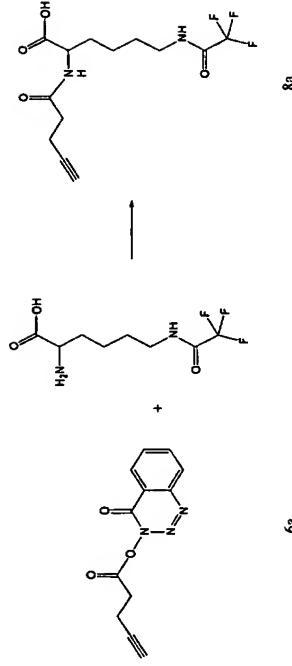
Phosphor oxy chloride (6.0  $\mu$ l, 0.059 mmol) was added to a cooled solution (0 °C) of 7a (30 mg, 0.054 mmol) in trimethyl phosphate (1 mL). The mixture was stirred for 1h. A solution of bis-*n*-tributylammonium pyrophosphate (77 mg, 0.16 mmol) in DMF (1 mL) and tributylamine (40  $\mu$ l, 0.16 mmol) were added. Water (2 mL) was added. pH of the solution was measured to be neutral. The solution was stirred at room temperature for 3 h and at 5 °C overnight. A small amount of compound VII (few mg) was obtained after purification by RP-HPLC (eluent: 100mM triethylammonium acetate → 20% acetonitrile in 100mM triethylammonium acetate). 7a (18 mg) was recovered.

Immediately prior to or subsequent to incorporation the protective di-*tert*-butyl ester groups may be cleaved to obtain the corresponding free carboxylic acid.

**Examples 34 and 35: Preparation of the mononucleotide building block**

Example 34: Preparation of 2-Pent-4-ynoylaminoo-6-(2,2-trifluoroacetylamino)-hexanoic acid, (8a)

The triethylammonium salt of compound VIII (11 mg) was obtained from compound 8a (50 mg, 0.15 mmol) and 5-iodo-5'-O-triphosphate-2'-deoxyuridine (50 mg, 0.06 mmol) using the procedure described for the synthesis of compound VI.



Compound 6a (250 mg, 1.0 mmol) was added to a solution of *N*-ε-trifluoroacetyl-L-lysine (Novabiochem, 04-125245) (250 mg, 1.0 mmol) in DMF (3 mL). Ethyl diisopropylamine (0.2 mL, 1.2 mmol) was added. The solution was stirred at room temperature overnight and worked-up by RP-HPLC (eluent: water → methanol). Yield: 50 mg, 15%.  $^1\text{H-NMR}$  ( $\text{D}_2\text{O}$ ): δ 4.4 (1H, CH), 3.4 (2H,  $\text{CH}_2$ ), 2.5 (4H,  $2 \times \text{CH}_2$ ), 2.3 (1H, CH), 1.9 (1H,  $\text{CH}_2$ ), 1.8 (1H,  $\text{CH}_2$ ), 1.6 (2H,  $\text{CH}_2$ ), 1.5 (2H,  $\text{CH}_2$ ).

Example 35: Preparation of 2-[5-[1-(4-Hydroxy-5-(O-triphosphatohydroxymethyl)-tetrahydrofuran-2-yl)-2,4-dioxo-1,2,3,4-tetrahydro-pyrimidin-5-yl]-pent-4-ynylamino]-9-(2,2,2-trifluoro-acetylamino)-hexanoic acid (Compound VIII)

三

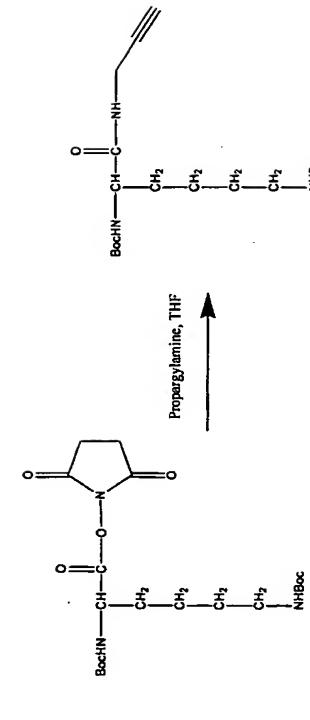
10

2

८

### Examples 36 to 40: Preparation of the monomer/oxide building block (X)

**Example 36: Preparation of di-Boc-Lysin-propargyl amide (compound 9a)**

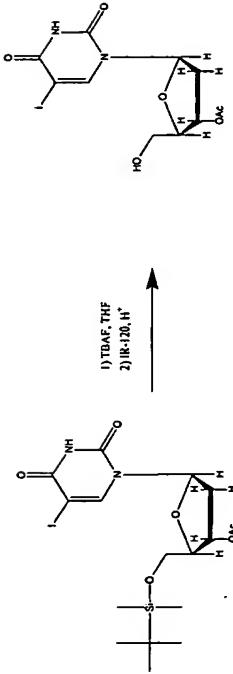


Boc-Lys-(Boc)-OSu (Novabiochem 04-12-0017, 0.887 g, 2 mmol) was dissolved in THF (10 ml). Propargylamine (0.412 ml, 6 mmol) was added and the solution stirred for 2 h. TLC (ethylacetate:heptan 1:1) showed only one product. Dichloromethane (20 ml) was added and the mixture was washed successively with citric acid (1M, 10 ml) and saturated sodium hydrogen carbonate (10 ml). The organic phase was dried with magnesium sulphate filtered and evaporated to give compound 9a (0.730 g) as a colourless syrup.

15

<sup>1</sup>H-NMR: δ 6.55 (1H, NH), 5.15 (1H, NH), 4.6 (1H, CH-NH), 4.05 (2H, CH-C-CH<sub>2</sub>-N), 3.75 (1H, NH), 3.1 (2H, CH<sub>2</sub>-NH) 2.25 (1H, CH-C-CH<sub>2</sub>), 1.9-1.3 (6H, 3 x CH<sub>2</sub>), 1.4 (18H, 6 x CH<sub>3</sub>).

**Example 37: Preparation of 5-iodo-3'-O-acetyl-5'-O-TBDMS-2'-deoxyuridine (compound 9b)**  $C_{17}H_{27}IN_2O_6Si$  Mw 510.40



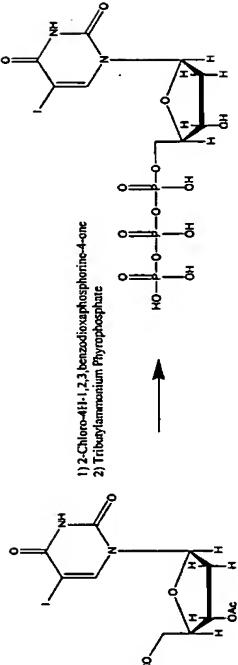
5-*Iodo-2'-deoxyuridine* (**Sigma I-7125**, 2.50 g, 7.08 mmol) and *imidazole* (0.961 g, 14.12 mmol) was dissolved in DMF (10 ml). Cooled to 0 °C and a solution of *TBDMSCl* (*t*-butyl-dimethyl-chloride, 1.12 g, 7.41 mmol) in dichloromethane (5.0 ml) was run in over 20 minutes. Stirring was continued at room temperature for 18 h, and the mixture was evaporated. The crude mono silylated nucleoside was dissolved in pyridine (40 ml) and cooled to 0 °C. *Acetic anhydride* (4.0 ml, 42.3 mmol) was added over 30 minutes and stirring was continued for 18 h at room temperature. The reaction mixture was evaporated and dissolved in dichloromethane (20 ml) and *citric acid* (2M, 20 ml) was added. The aqueous phase was back extracted with dichloromethane (2 x 20 ml). The combined organic phases were washed with saturated sodium bicarbonate (20 ml), dried with sodium sulphate and evaporated (5.85 g). Recrystallisation form ethyl acetate/EtOH gave **9b** (2.54, g) pure for synthesis **TLC** (Ethyl acetate). Further recrystallisation furnished an analytical pure sample mp. 172.4-173.1 °C.

Example 38: Preparation of 5-*lodo*-3'-O-acetyl-2'-deoxyuridine (compound 9c)  $C_{11}H_{13}IN_2O_6$  Mw 396.14

20

5-ido-3'-O-acetyl-5'-O-TBDMS-2'-deoxyuridine (compound 9b) (2.54 g, 4.98 mmol) as dissolved in THF (25 ml), tetra butyl ammonium fluoride trihydrat (TBAF, 3.2 g, 10.1 mmol) was added and stirred for 18 h at room temperature. The reaction mixture was added water (25 ml) stirred for 1 h. Ion exchange resin IR-120 H<sup>+</sup> (26 ml) was then added and stirring was continued for 1 h. The solution was filtered and reduced to approximately 10 ml in vacuo. Crystals were collected and dried in vacuo (1.296g)

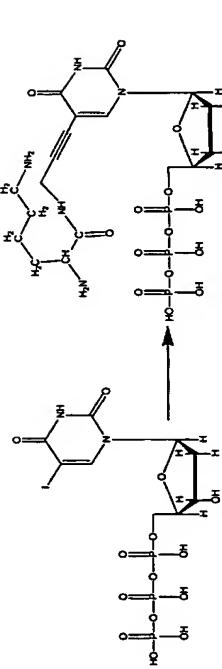
10 Example 39: Preparation of 5-iodo-5'-O-triphosphate-2'-deoxyuridine, triethylammonium salt (compound 9d) C<sub>9</sub>H<sub>14</sub>[N<sub>2</sub>O<sub>4</sub>P<sub>3</sub>] + n-N(CH<sub>2</sub>CH<sub>3</sub>)<sub>3</sub> Mw 897.61 for n=3.



15 5-ido-3'-O-acetyl-2'-deoxyuridine (compound 9c) (2.54 g, 4.98 mmol) as dissolved in pyridine (3.2 ml) and dioxane (10 ml). A solution of 2-chloro-4H-1,3,2-benzodioxaphosphorin-4-one in dioxane (3.60 ml, 1 M, 3.60 mmol) was added under stirring. The reaction mixture was stirred for 10 minutes at room temperature followed by simultaneous addition of bis(tri-n-butylammonium) pyrophosphate in DMF (9.81 ml, 0.5 M, 4.91 mmol) and tri-n-butylamine (3.12 ml, 13.1 mmol). Stirring was continued for 10 minutes and the intermediate was oxidized by adding an iodine solution (90 ml, 1% w/v in pyridine/water (98/2, v/v)) until permanent iodine colour. The reaction mixture was left for 15 minutes and then decolourised with sodium thiosulfate (5% aqueous solution, w/v). The reaction mixture was evaporated to yellow oil. The oil was stirred in water (20 ml) for 30 minutes and concentrated aqueous ammonia (100 ml, 25%) was added. This mixture was stirred for 1.5 hour at room temperature and then evaporated to an oil of the crude triphosphate product. The

crude material was purified using a DEAE Sephadex A25 column (approximately 100 ml) eluted with a linear gradient of triethyl- ammonium hydrogen carbonate [TEAB] from 0.05 M to 1.0 M (pH approximately 7.0 – 7.5); flow 8 ml/fraction/15 minutes. The positive fractions were identified by RP18 HPLC eluting with a gradient from 10 mM TEAA (triethylammonium acetate) in water to 10 mM TEAA 20% water in acetonitrile. The appropriate fractions were pooled and evaporated. Yield approximately 1042 mg.

10 Example 40: Preparation of 5-(Lysin-propargyl amide)-5'-triphosphate-2'-deoxyuridine, triethylammonium salt (compound IX) C<sub>18</sub>H<sub>30</sub>N<sub>5</sub>O<sub>15</sub>P<sub>3</sub> + n-N(CH<sub>2</sub>CH<sub>3</sub>)<sub>3</sub> Mw 952.95 for n=3

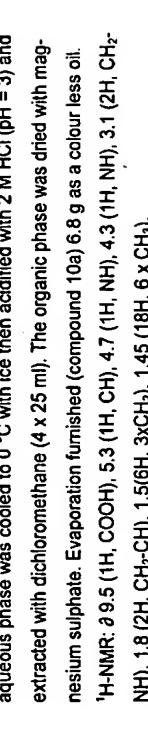


15 5-ido-3'-O-acetyl-5'-triphosphate-2'-deoxyuridine, triethylammonium salt (compound 9d) (0.0087 g, 9.17 μmol) was dissolved in water (100 μl). Air was replaced carefully with argon. Di-Boc-Lysin-propargyl amide (compound 9a) (18.6 mg, 48.5 μmol) dissolved in dioxane (100 μl), triethylamine (2.7 μl, 19.4 μl), Pd((PPh<sub>3</sub>)<sub>4</sub> C<sub>6</sub>H<sub>5</sub>SO<sub>3</sub>Na<sup>+</sup>)(H<sub>2</sub>O)<sub>4</sub> (compound 9d) (5 mg, 4.4 μmol) and copper (I) iodide (0.4 μl, 2.1 μmol) were added in the given order. The reaction mixture was stirred for 18 h at room temperature in an inert atmosphere then evaporated. The crude material was treated with aqueous hydrochloric acid (0.2 M, 1 ml) for 15 minutes at 30 °C. (compound IX) was obtained by HPLC C<sub>18</sub> 10 mM TEAA (triethylammonium acetate) in water to 10 mM TEAA 20% water in acetonitrile. Appropriate fractions were desalting using gel filtration (pharmacia G-10, 0.7 ml).

**Examples 41 to 46: Preparation of the mononucleotide building block (X)**

**Example 41: Preparation of Boc-Lys-1(Boc)-OH (compound 10a)  $C_{16}H_{30}N_2O_6$  Mw 346.42**

5 Lysine (Novabiochem 04-10-0024; 3.65 g, 20 mmol) was dissolved in sodium hydroxide (2 M, 40 ml), added dioxane (60 ml) and di-tert-butyl dicarbonate (8.73 g, 40 mmol) in the given order. The mixture was stirred for 1.75 h at 60 °C. Water (50 ml) was added and the solution was washed with dichloromethane (4 x 25 ml). The aqueous phase was cooled to 0 °C with ice then acidified with 2 M HCl (pH = 3) and extracted with dichloromethane (4 x 25 ml). The organic phase was dried with magnesium sulphate. Evaporation furnished (compound 10a) 6.8 g as a colour less oil.  $^1H$ -NMR:  $\delta$  9.5 (1H, COOH), 5.3 (1H, CH), 4.7 (1H, NH), 4.3 (1H, NH), 3.1 (2H, CH<sub>2</sub> NH), 1.8 (2H,  $\underline{CH_2}CH$ ), 1.5 (6H, 3xCH<sub>3</sub>), 1.45 (18H, 6 x CH<sub>3</sub>).

**Example 42: Preparation of di-Boc-Lysin-propargyl ester (compound 10b)  $C_{18}H_{32}N_2O_6$  Mw 384.47**

15

15

Boc-Lys-(Boc)-OH (compound 10a) (3.46 g, 10 mmol) was dissolved in THF (25 ml). At 0 °C a solution of dicyclohexylcarbodiimide (2.02 g, 9.0 mmol) in THF (25 ml) and triethylamine (1.39 ml) were added in the given order. The mixture was allowed to warm up to room temperature and stirred for 18 h. The resulting suspension was

20

SUBSTITUTE SHEET (RULE 26)

53

filtered and evaporated. The oil 5.45 g was pre-purified by column chromatography

Heptan: Ethylacetate 3:1.

Pure 10b was achieved by HPLC- C<sub>18</sub> 10% MeOH: 90% H<sub>2</sub>O → 100% MeOH  
 $^1H$ -NMR:  $\delta$  5.1 (1H, NH), 4.75 (2H, CH<sub>2</sub>-CH<sub>2</sub>O), 4.6 (1H, NH), 4.35 (1H,  $\underline{CH_2}NH$ ), 3.1 (2H,  $\underline{CH_2}NH$ ) 2.5 (1H,  $\underline{CH_2}CH_2$ ), 1.9-1.4 (6H, 3 x CH<sub>2</sub>), 1.5 (18H, 6 x CH<sub>3</sub>).

**Example 43: Preparation of 5-Iodo-3',5'-di-O-TBDMs-2'-deoxycytidine (compound 10c)  $C_{21}H_{40}N_4O_4S_2$  Mw 581.64**

10

10

5-Iodo-2-deoxy-Cytidine (Sigma I-7000, 0.353 g, 1 mmol) was dissolved in DMF (4 ml), added t-Butyl-dimethyl silyl chloride (TBDMSCl, 0.332 g, 2.2 mmol) and Imidazole (0.204 g, 3 mmol). The solution was stirred for 15 h at 50 °C followed by evaporation. Dichloromethane (25 ml) and citric acid (2M, 10 ml) was added to the dry mixture. The aqueous phase was back extracted with dichloromethane (2 x 10 ml). The combined organic phases were washed with saturated sodium bicarbonate (15 ml), dried with sodium sulphate and evaporated. Compound 10 c (0.405 g) was obtained by recrystallisation from EtOH/Ethylacetate.

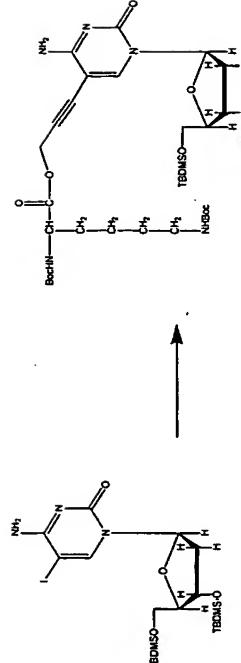
$^1H$ -NMR:  $\delta$  8.1 (1H, H-1'), 6.25 (1H, H-4'), 4.35 (1H, H-4), 3.0 (1H, H-5'), 3.75 (1H, H-5), 2.5 (1H, H-2'), 1.95 (1H, H-2), 1.85 (2H, NH), 0.95 (9H, 3 x CH<sub>3</sub>), 0.9 (9H, 3 x CH<sub>3</sub>), 0.15 (6H, 2 x CH<sub>3</sub>), 0.1 (6H, 2 x CH<sub>3</sub>).

**Preparation of 5-(di-Boc-Lysin-propargyl ester)-3',5'-di-O-TBDMs-2'-deoxycytidine (compound 10d)  $C_{40}H_{71}IN_5O_{10}S_2$  Mw 838.19**

25

SUBSTITUTE SHEET (RULE 26)

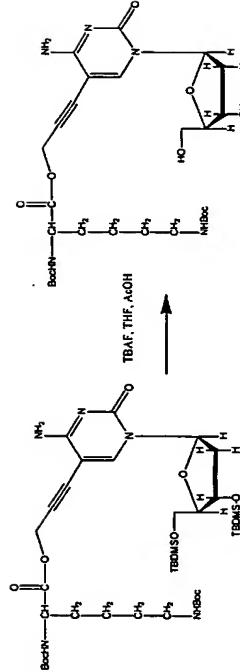
54



Compound 10c (0.116 g, 0.2 mmol) was dissolved in dichloromethane (10 ml). Air was replaced carefully with argon. Di-Boc-Lysin-propargyl ester (compound 10b) (0.232, 0.6 mmol), triethylamine (0.038 ml, 0.6 mmol), di-chloro-bis-triphenylphosphine-palladium II (0.0074 g, 0.01 mmol) and copper (I) iodide (0.0038 g, 0.02 mmol) were added in the given order. The reaction mixture was stirred for 15 h at room temperature in an inert atmosphere. The reaction mixture was evaporated re-dissolved in MeOH/H<sub>2</sub>O 1:1 ml and purified using HPLC-C<sub>18</sub> 45% H<sub>2</sub>O:55% MeCN → 100% MeCN.

<sup>1</sup>H-NMR: δ (H-NMR): 3.82 (1H, H-6), 6.25 (1H, H-1), 5.15 (1H, NH), 4.9 (2H, C=CH<sub>2</sub>-O), 4.6 (1H, NH), 4.4 (1H, H-4'), 4.3 (1H, CH<sub>2</sub>-NH), 4.0 (1H, H-4'), 3.9 (1H, H-5'), 3.75 (1H, H-5), 2.5 (1H, H-2), 3.1 (2H, CH<sub>2</sub>-NH), 1.95 (1H, H-2'), 1.9-1.4 (6H, 3 x CH<sub>3</sub>), 1.85 (2H, NH), 1.5 (18H, 6 x CH<sub>3</sub>), 0.95 (9H, 3 x CH<sub>3</sub>), 0.15 (6H, 2 x CH<sub>3</sub>), 0.1 (6H, 2 x CH<sub>3</sub>).

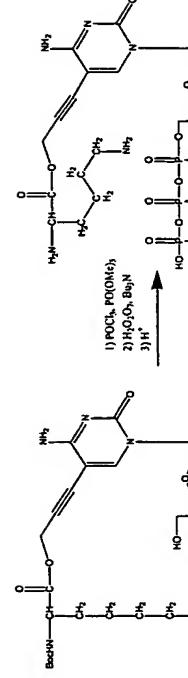
**Example 44: Preparation of 5-(di-Boc-Lysin-propargyl ester)-2'-deoxycytidine (compound 10e) C<sub>28</sub>H<sub>41</sub>N<sub>5</sub>O<sub>10</sub> Mw 609.67**



Compound 10d (0.0246 g, 0.029 mmol) was dissolved in THF (1 ml) and successively added acetic acid (0.0165 ml, 0.288 mmol) and tetra n-butyl ammonium fluoride tri-hydrate (0.0454 g, 0.144 mmol). The reaction mixture was stirred for 18 h at room temperature and afterwards evaporated. Re-dissolved in dichloromethane and purified on silica (1 x 18 cm). Dichloromethane/MeOH 8/2. Fractions which gave UV absorbance on TLC were pooled and evaporated giving 10e (0.0128 g) as a colourless oil.

**Example 45: Preparation of 5-(Lysin-propargyl ester)-5'-triphosphate-2'-deoxycytidine C<sub>18</sub>H<sub>30</sub>N<sub>5</sub>O<sub>5</sub>P<sub>3</sub> Mw 649.38**

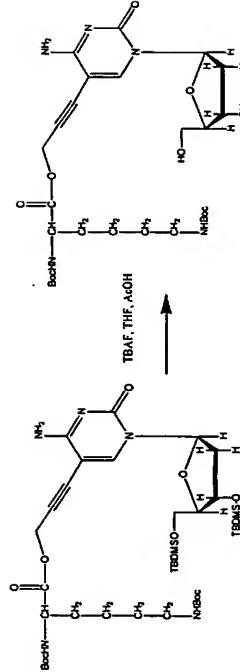
10

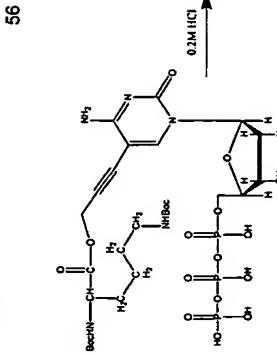


Compound 10d (0.0128 g, 0.021 mmol) was dissolved in trimethylphosphate (0.150 ml) and cooled to 0 °C. Phosphotrichloride in trimethylphosphate (1M, 0.0246 ml) was added slowly in order not to raise the temperature. Stirring was continued for 2 h at 0 °C and the temperature was allowed to rise to ambient. Pyrophosphate in DMF (0.5 M, 0.1025 ml, 0.051 mmol) and tri-n-butyl amine in DMF (1M, 0.0122 ml, 0.051 mmol) were added simultaneous. Stirring was continued for 15 minutes at room temperature and TEAB[triethyl ammonium bicarbonate, 1M, pH = 7.3, 0.50ml] was added. Stirring was continued for 3 h then evaporated.

20

**Example 46: Preparation of compound X**

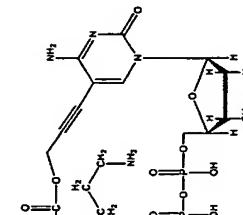




The crude material was treated with aqueous hydrochloric acid (0.2 M, 1 ml) for 15 minutes at 30 °C. Compound X was obtained by HPLC C<sub>18</sub> 10 mM TEAA (triethylammonium acetate) in water to 10 mM TEAA 20% water in acetonitrile. Appropriate fractions were desalting using gel filtration (pharmacia G-10, 0.7 ml)

**Example 47: Polymerase incorporation of different nucleotide derivatives.**

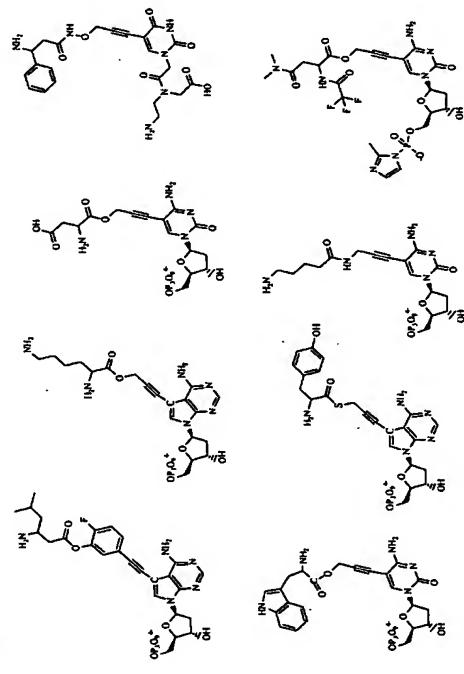
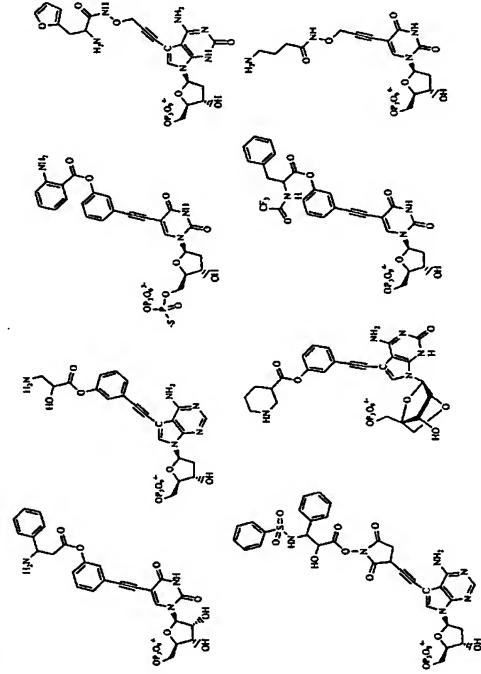
Different extension primers were 5'-labeled with <sup>32</sup>P using T4 polynucleotide kinase using standard protocol (Promega, cat# 4103). These extension primers was annealed to a template primer using 0.1 and 3 pmol respectively in an extension buffer (20 mM H<sub>2</sub>PO<sub>4</sub>, 40 mM KCl, 8 mM MgCl<sub>2</sub>, pH 7.4, 10 mM DTT) by heating to 80 °C for 2 min. and then slowly cooling to about 20 °C. The wild type nucleotide or nucleotide derivatives was then added (about 100 μM) and incorporated using 5 units AMV Reverse Transcriptase (Promega, part# 9PIM510) at 30 °C for 1 hour. The samples were mixed with formamide dye and run on a 10% urea polyacrylamide gel electrophoresis. The gel was developed using autoradiography (Kodak, BioMax film). The incorporation can be identified by the different mobility shift for the nucleotide derivatives compared to the wild type nucleotide. Figure 1 shows incorporation of various nucleotide derivatives. In lane 1-5 the extension primer 5'-GGT ACT GGC ATC GGT-3' was used together with the template primer 5'-GCT GTC TGC AAG TGA TAA CCG ATG CCA GTA GC-3'; in lane 6-11 extension primer 5'-GCT ACT GGC ATC GGT-3' was used together with the template primer 5'-GCT GTC TGC AAG TGA TGA CCG ATG CCA GTA GC-3', and in lane 12-15 the extension primer 5'-GCT ACT GGC ATC GGT-3' was used together with the template primer 5'-GCT GTC TGC AAG TGA CCG AAC CGT AAC CGA TGC CAG TAG C-3'. Lane 1, dATP; lane 2, not relevant; lane 3, Compound IX; lane 4, Compound I; lane 5, Compound II; lane 6, no nucleotide; lane 7, dCTP; lane 8, Compound VII; lane 9, Compound X; lane



10, Compound IV; lane 11, Compound III; lane 12, no nucleotide; lane 13, dTTP; lane 14, dTTP and dATP; lane 15, dTTP and Compound X. These results illustrate the possibility to incorporate a variety of nucleotide derivatives of dATP, dTTP and dCTP using different linkers and functional entities. Other polymerases such as Taq, M-MLV and HIV have also been tested with positive results.

5

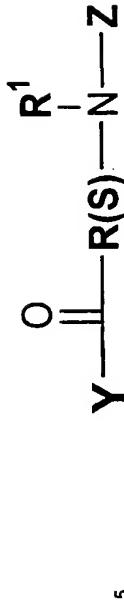
The compounds shown in chart 4 may be synthesised by the methods described above.



5 Chart 4 Building blocks for library preparation

## Claims

1. A Nucleoside derivative having the general formula:



Wherein Y is a group  $\text{---} \text{X} \text{---} \text{R}^2 \text{---} \text{C} \equiv \text{C} \text{---} \text{N} \text{---}$

Wherein

X is a hetero atom selected from the group O, S, Se or a group  $\text{NR}^4$ , wherein  $\text{R}^4$  is hydrogen or an optionally substituted linear or branched  $\text{C}_{1-6}$  alkyl or  $\text{C}_{2-6}$  alkenyl.

$\text{R}^2$  is selected from the group consisting of  $\text{C}_{1-6}$  alkyl,  $\text{C}_{2-6}$  alkylene,  $\text{C}_{2-6}$  alkylenyl,  $\text{C}_{2-6}$  cycloalkyl, heterocycloalkyl, - $\text{CH}_2\text{O}$ , aryl or heteroaryl, wherein each of the groups  $\text{R}^2$  are substituted with 0-3  $\text{R}^8$  groups independently selected from =O, =S, =F, -Cl, -Br, -I, -OCH<sub>3</sub>, -NO<sub>2</sub> or  $\text{C}_{1-6}$  alkyl, and

N is a nucleoside analogue consisting of a nucleobase and a backbone unit;

or Y is  $\text{---} \text{OR}^3$ , wherein  $\text{R}^3$  is H or an acid protective group.

$\text{R}(\text{S})$  is a  $\text{C}_{1-4}$  alkylen,  $\text{C}_{3-10}$  cycloalkyl, aryl, heterocycloalkyl or heteroaryl substituted by sidechains S, wherein n is an integer of 0 to 4

10  $\text{R}^1$  is H,  $\text{C}_{1-6}$  alkyl substituted with 0-3  $\text{R}^6$  where  $\text{R}^6$  is independently selected from =O, Cl, Br, -CN, -OR<sup>6</sup>, -SR<sup>6</sup>, -NR<sup>6</sup>R<sup>7</sup>, -CONR<sup>6</sup>R<sup>7</sup>, -SO<sub>2</sub>NR<sup>6</sup>R<sup>7</sup> or a  $\text{C}_{1-6}$  alkylene group forming a ringstructure with S

15  $\text{R}^6$  and  $\text{R}^7$  are independently selected from H,  $\text{C}_{1-6}$  linear alkyl,  $\text{C}_{1-6}$  branched alkyl,  $\text{C}_{1-6}$  cycloalkyl, aryl, heteroaryl, or hetero aralkyl.

20  $\text{S}$  is  $\text{C}_{1-6}$  linear alkyl,  $\text{C}_{3-6}$  branched alkyl,  $\text{C}_{3-6}$  cycloalkyl, aryl, heteroaryl, aralkyl, hetero aralkyl substituted with 0-3  $\text{R}^3$  where  $\text{R}^3$  is independently selected from =O, Cl, Br, -CN, -OR<sup>6</sup>, -SR<sup>6</sup>, -NR<sup>6</sup>R<sup>7</sup>, -CONR<sup>6</sup>R<sup>7</sup>, -SO<sub>2</sub>NR<sup>6</sup>R<sup>7</sup>.

**Z** is H, an amino protective group or a group  $\text{---C}=\text{R}^2\text{---C}\equiv\text{C---Ns}$  with the proviso, that when **Y** is not  $\text{---X---R}^2\text{---C}\equiv\text{C---Ns}$ , **Z** is  $\text{---C}=\text{R}^2\text{---C}\equiv\text{C---Ns}$

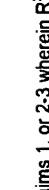
2. A compound according to claim 1 wherein the alkynylen linker is connected to the nucleobase of a nucleoside analogue.

3. A compound according to claim 1 wherein the alkynylen linker is connected to the nucleobase of a nucleoside analogue in the 7 position of the bicyclic purine nucleobases and the 5 position of the monocyclic pyrimidine bases.

4. A compound according to any of the claims 1, or 2-3 wherein  $\text{R}^2$  is selected from the group consisting of  $\text{C}_{1-6}$  alkylén,  $\text{C}_{2-6}$  alkenylen,  $\text{C}_{2-6}$  alkynylén, heterocycloalkylen,  $\text{---CH}_2\text{O}$ , aryen or heteroarylen, wherein each of the groups  $\text{R}^2$  are substituted with 0-3  $\text{R}^6$  groups independently selected from =O, -F, -Cl, -Br, -NO<sub>2</sub>,  $\text{C}_{1-4}$  alkyl.

5. A compound according to any of the claims 1, or 2-3 wherein  $\text{R}^2$  is selected from the group consisting of  $\text{C}_{1-6}$  alkylén,  $\text{C}_{2-6}$  alkenylen,  $\text{C}_{2-6}$  alkynylén, heterocycloalkylen,  $\text{---CH}_2\text{O}$ , aryen or heteroarylen, wherein each of the groups  $\text{R}^2$  are substituted with 0-2  $\text{R}^6$  groups independently selected from =O, -F, -NO<sub>2</sub>,  $\text{C}_{1-6}$  alkyl.

6. A compound according to any of the claims 1, or 2-3 wherein  $\text{R}^2$  is selected from the group consisting of  $\text{---CH}_2\text{r}$ ,  $\text{---CH}_2\text{CH}_2$ , ,  $\text{---CH}_2\text{O}$ , or aryen wherein each of the groups  $\text{R}^2$  are substituted with 0-2  $\text{R}^6$  groups independently selected from =O, -F, -NO<sub>2</sub>,  $\text{C}_{1-6}$  alkyl.

7. A compound according to any of the claims 1, or 2-3 wherein  $\text{R}^2$  is selected from the group consisting of  $\text{---CH}_2\text{r}$ ,  $\text{---CH}_2\text{CH}_2$ , ,  $\text{---CH}_2\text{O}$ , or aryen.

8. A compound according to any of the claims 1, or 2-3 wherein  $\text{R}^2$  is selected from the group consisting of  $\text{---CH}_2\text{r}$ ,  $\text{---CH}_2\text{CH}_2$ ,  or aryen.

9. A compound according to any of the claims 1, 2-3 or 4-8 wherein **X** is O

5 10. A compound according to any of the claims 1, 2-3 or 4-8 wherein **X** is S  
11. A compound according to any of the claims 1, 2-3 or 4-8 wherein **X** is NR<sup>4</sup>

12. A compound according to any of the claims 1, 2-3 or 4-8 wherein **X** is NR<sup>4</sup> and R<sup>4</sup> is H or -CH<sub>3</sub>

10 13. A compound according to any of the claims 1, 2-3 or 4-8 wherein **X** is NH

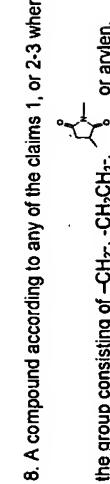
14. A compound according to any of the claims 1, 2-3, 4-8 or 9, 10 or 11-13 wherein R(S) is a  $\text{C}_{1-4}$  alkylene,  $\text{C}_{3-10}$  cycloalkylene, aryl, heterocycloalkyl or heteroaryl substituted by n sidechains S, wherein n is an integer of 0 to 3

15 16. A compound according to any of the claims 1, 2-3, 4-8 or 9, 10 or 11-13 wherein R(S) is a  $\text{C}_{1-4}$  alkylene, aryl or heteroaryl substituted by n sidechains S, wherein n is an integer of 0 to 3

20 17. A compound according to any of the claims 1, 2-3, 4-8 or 9, 10 or 11-13 wherein R(S) is a  $\text{C}_{1-4}$  alkylene substituted by n sidechains S, wherein n is an integer of 0 to 3

25 18. A compound according to any of the claims 1, 2-3, 4-8 or 9, 10 or 11-13 wherein R(S) is a  $\text{C}_{1-2}$  alkylene substituted by n sidechains S, wherein n is an integer of 0 to 2

30 19. A compound according to any of the claims 1, 2-3, 4-8 or 9, 10 or 11-13 wherein R(S) is a  $\text{C}_{1-2}$  alkylene substituted by n sidechains S, wherein n is an integer of 0 to 1



62

20. A compound according to any of the claims 1, 2-3, 4-8, 9, 10, 11-13 or 14-19 wherein S is C<sub>1-4</sub> linear alkyl, C<sub>3-6</sub> branched alkyl, C<sub>3-6</sub> cycloalkyl, aryl, heteroaryl, aralkyl, hetero aralkyl substituted with 0-3 R<sup>5</sup> where R<sup>5</sup> is independently selected from =O, Cl, Br, -CN, -OR<sup>6</sup>, -SR<sup>6</sup>, -NR<sup>6</sup>R<sup>7</sup>, -COOR<sup>6</sup>, -CONRR<sup>7</sup>, -SO<sub>2</sub>NR<sup>6</sup>R<sup>7</sup> where R<sup>6</sup> and R<sup>7</sup> are independently selected from H, C<sub>1-3</sub> linear alkyl, C<sub>3-6</sub> cycloalkyl, aryl, heteroaryl, aralkyl, or hetero aralkyl.

21. A compound according to any of the claims 1, 2-3, 4-8, 9, 10, 11-13 or 14-19 wherein S is C<sub>1-4</sub> linear alkyl, C<sub>3-6</sub> branched alkyl, C<sub>3-6</sub> cycloalkyl, aryl, heteroaryl, aralkyl, hetero aralkyl substituted with 0-2 R<sup>5</sup> where R<sup>5</sup> is independently selected from =O, Cl, -CN, -OR<sup>6</sup>, -SR<sup>6</sup>, -NR<sup>6</sup>R<sup>7</sup>, -COOR<sup>6</sup>, -CONRR<sup>7</sup>, -SO<sub>2</sub>NR<sup>6</sup>R<sup>7</sup> where R<sup>6</sup> and R<sup>7</sup> are independently selected from H, C<sub>1-3</sub> linear alkyl, aryl, heteroaryl, aralkyl, or hetero aralkyl.

22. A compound according to any of the claims 1, 2-3, 4-8, 9, 10, 11-13 or 14-19 wherein S is C<sub>1-4</sub> linear alkyl, C<sub>3-6</sub> branched alkyl, C<sub>3-6</sub> cycloalkyl, aryl, heteroaryl, aralkyl, hetero aralkyl substituted with 0-2 R<sup>5</sup> where R<sup>5</sup> is independently selected from =O, Cl, -CN, -OR<sup>6</sup>, -SR<sup>6</sup>, -NR<sup>6</sup>R<sup>7</sup>, -COOR<sup>6</sup>, -CONRR<sup>7</sup>, -SO<sub>2</sub>NR<sup>6</sup>R<sup>7</sup> where R<sup>6</sup> and R<sup>7</sup> are independently selected from H and C<sub>1-3</sub> linear alkyl

23. A compound according to any of the claims 1, 2-3, 4-8, 9, 10, 11-13 or 14-19 wherein S is C<sub>1-4</sub> linear alkyl, C<sub>3-6</sub> branched alkyl, C<sub>3-6</sub> cycloalkyl, aryl, heteroaryl, aralkyl, hetero aralkyl substituted with 0-1 R<sup>5</sup> where R<sup>5</sup> is selected from =O, Cl, -CN, -OR<sup>6</sup>, -SR<sup>6</sup>, -NR<sup>6</sup>R<sup>7</sup>, -COOR<sup>6</sup>, -CONRR<sup>7</sup>, -SO<sub>2</sub>NR<sup>6</sup>R<sup>7</sup> where R<sup>6</sup> and R<sup>7</sup> are independently selected from H and C<sub>1-3</sub> linear alkyl

24. A compound according to any of the claims 1, 2-3, 4-8, 9, 10, 11-13 or 14-19 wherein S is C<sub>1-4</sub> linear alkyl or aryl substituted with 0-1 R<sup>5</sup> where R<sup>5</sup> is selected from =O, Cl, -CN, -OR<sup>6</sup>, -SR<sup>6</sup>, -NR<sup>6</sup>R<sup>7</sup>, -COOR<sup>6</sup>, -CONRR<sup>7</sup>, -SO<sub>2</sub>NR<sup>6</sup>R<sup>7</sup> where R<sup>6</sup> and R<sup>7</sup> are independently selected from H and C<sub>1-3</sub> linear alkyl

25. A compound according to any of the claims 1, 2-3, 4-8, 9, 10, 11-13 or 14-19 wherein S is C<sub>1-4</sub> linear alkyl or aryl.

35

26. A compound according to any of the claims 1, 2-3, 4-8, 9, 10, 11-13, 14-19 or 20-25 wherein R<sup>1</sup> is H, C<sub>1-6</sub> alkyl substituted with 0-1 R<sup>9</sup> where R<sup>9</sup> is independently selected from =O, Cl, Br, -CN, -OR<sup>6</sup>, -SR<sup>6</sup>, -NR<sup>6</sup>R<sup>7</sup>, -COOR<sup>6</sup>, -CONRR<sup>7</sup>, -SO<sub>2</sub>NR<sup>6</sup>R<sup>7</sup> where R<sup>6</sup> and R<sup>7</sup> are independently selected from H, C<sub>1-6</sub> linear alkyl, C<sub>1-6</sub> branched alkyl, C<sub>1-6</sub> cycloalkyl, aralkyl, or hetero aralkyl or a C<sub>1-6</sub> alkylene group forming a ringstructure with S.

27. A compound according to any of the claims 1, 2-3, 4-8, 9, 10, 11-13, 14-19 or 20-25 wherein R<sup>1</sup> is H, C<sub>1-6</sub> alkyl or a C<sub>1-6</sub> alkylene group forming a ringstructure with S

10

28. A compound according to any of the claims 1, 2-3, 4-8, 9, 10, 11-13, 14-19 or 20-25 wherein R<sup>1</sup> is H or a C<sub>1-6</sub> alkylene group forming a ringstructure with S.

29. A compound according to any of the claims 1, 2-3, 4-8, 9, 10, 11-13, 14-19 or 20-25 wherein R<sup>1</sup> is H or C<sub>1-6</sub> alkyl.

30. A compound according to any of the claims 1, 2-3, 4-8, 9, 10, 11-13, 14-19 or 20-25 wherein R<sup>1</sup> is H.

31. A compound according to any of the claims 1, 2-3, 4-8, 9, 10, 11-13, 14-19, 20-25 or 26-30 wherein Z is H, an amino protective group selected from the group of formyl, acetyl, trifluoroacetyl, benzoyl, tert-butoxycarbonyl, triphenylmethyl, benzyl, benzyloxycarbonyl or tosyl or a group  $\text{O}=\text{C}-\text{R}^2-\text{C}\equiv\text{C}-\text{Ns}$  with the proviso, that

$\text{O}=\text{C}-\text{R}^2-\text{C}\equiv\text{C}-\text{Ns}$  with the proviso, that when Y is not  $\text{O}=\text{C}-\text{R}^2-\text{C}\equiv\text{C}-\text{Ns}$

25 when Y is not  $\text{O}=\text{C}-\text{R}^2-\text{C}\equiv\text{C}-\text{Ns}$  Z is  $\text{O}=\text{C}-\text{R}^2-\text{C}\equiv\text{C}-\text{Ns}$

32. A compound according to any of the claims 1, 2-3, 4-8, 9, 10, 11-13, 14-19, 20-25 or 26-30 wherein Z is H, an amino protective group selected from the group of acetyl, trifluoroacetyl, tert-butoxycarbonyl or tosyl or a group  $\text{O}=\text{C}-\text{R}^2-\text{C}\equiv\text{C}-\text{Ns}$  with the proviso, that when Y is not  $\text{O}=\text{C}-\text{R}^2-\text{C}\equiv\text{C}-\text{Ns}$  Z is

30  $\text{O}=\text{C}-\text{R}^2-\text{C}\equiv\text{C}-\text{Ns}$



33. A compound according to any of the claims 1, 2-3, 4-8, 9, 10, 11-13, 14-19, 20-25, 26-30 or 31-32 wherein the nucleobase is uracil or cytosine modified in the 5 position or 7-adeazaadenine or 7-deazaguanidine modified in the 7 position.

5

34. A compound according to any of the claims 1, 2-3, 4-8, 9, 10, 11-13, 14-19, 20-25, 26-30, 31-32 or 33 wherein the backbone unit type is DNA, RNA, Oxy-LNA, Thio-LNA, Amino-LNA, Phosphothioate, 2'-O-methyl, PNA or Morpholino as described in chart 3.

10

35. A compound according to any of the claims 1, 2-3, 4-8, 9, 10, 11-13, 14-19, 20-25, 26-30, 31-32 or 33 wherein the backbone unit type is DNA, RNA, Oxy-LNA, PNA or Morpholino

15

36. A compound according to any of the claims 1, 2-3, 4-8, 9, 10, 11-13, 14-19, 20-25, 26-30, 31-32 or 33 wherein the backbone unit type is DNA, PNA or Oxy-LNA

37. A compound according to any of the claims 1, 2-3, 4-8, 9, 10, 11-13, 14-19, 20-25, 26-30, 31-32 or 33 wherein the backbone unit type is DNA

20

38. A compound according to any of the claims 1, 2-3, 4-8, 9, 10, 11-13, 14-19, 20-25, 26-30, 31-32 or 33 wherein the backbone unit type is Oxy-LNA

39. A compound according to any of the claims 1, 2-3, 4-8, 9, 10, 11-13, 14-19, 20-25, 26-30, 31-32 or 33 wherein the backbone unit type is PNA

25

40. A compound according to any of the claims 1, 2-3, 4-8, 9, 10, 11-13, 14-19, 20-25, 26-30, 31-32, 33 or 34-39 wherein more nucleoside analogues are connected via their backbone structures forming di-, tri- or oligomeric nucleoside analogues as building blocks

30

41. A compound according to any of the claims 1, 2-3, 4-8, 9, 10, 11-13, 14-19, 20-25, 26-30, 31-32, 33, 34-39 or 40 wherein Y is  $-X-R^2-C\equiv C-Ns$  or  $-OR^3$  wherein  $R^3$  is selected from the group H, C<sub>1-3</sub> alkyl, allyl, benzyl, *tert*-butyl or triphenylmethyl.

35



## INTERNATIONAL SEARCH REPORT

International Application No.  
PCT/UK 02/00420

### Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  **Claims No.**, because they relate to subject matter not required to be searched by this Authority, namely:

2.  **Claims No.**, because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:  
see FURTHER INFORMATION sheet PCT/ISA/210

3.  **Claims No.**, because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a),

### Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this International application, as follows:

1.  As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not make payment of any additional fee.
3.  As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claim Nos.:
4.  No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claim Nos.:

Remark on Patent

The additional search fees were accompanied by the applicant's protest.  
 No protest accompanied the payment of additional search fees.

International Application No. PCT/UK 02/00420

### FURTHER INFORMATION CONTINUED FROM PCT/ISA/210

#### Continuation of Box I.2

Present claims 1-41 relate to an extremely large number of possible compounds. Support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT is to be found, however, for only a very small proportion of the compounds claimed. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Consequently, the search has been carried out for those parts of the claims which appear to be supported and disclosed, namely those parts relating to the compounds of the general formula in claim 1, wherein:

(1) the alkynylen linker is connected to the nucleobase of a nucleoside analogue in the 7 position of the bicyclic purine nucleobases and the 5 position of the monocyclic pyrimidine bases (as defined in claim 3), and

(2) R(S) is a C1-4 alkylene, and

(3) the backbone unit-type is DNA or RNA.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no International search report has been established need not be the subject of an International preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

**INTERNATIONAL SEARCH REPORT**

Information on patent family members\*

Patient document cited in search report		Publication date	Patient family member(s)	Publication date
WO 9737041	A 09-10-1997	US AU WO US	6194144 B1 2217597 A 9737041 A2 6238871 B1	27-02-2001 22-10-1997 09-10-1997 29-05-2001
WO 9421822	A 29-09-1994	AT AU AU CA DE EP JP WO US US US US	220114 T 687801 B2 6411694 A 2158642 A1 69438909 D1 0689610 A1 8507926 T 9421822 A1 6140053 A 5622284 A 5872003 A 5851765 A 6074823 A	15-07-2002 05-03-1998 11-10-1994 29-09-1994 08-08-2002 03-01-1996 27-08-1996 29-09-1994 31-10-2000 22-04-1997 16-02-1999 22-12-1998 13-06-2000
WO 9416101	A 21-07-1994	AU AU CA EP JP WO US US US US US US US US	694940 B2 599294 A 9137998 A 2153387 A1 0679196 A1 8509857 T 9416101 A2 5547835 A 5605798 A 5691141 A 6225450 B1 6238871 B1 6436635 B1 6194144 B1	06-08-1998 15-08-1994 14-01-1999 21-07-1994 02-11-1995 22-10-1996 21-07-1994 20-08-1996 25-02-1997 25-11-1997 01-05-2001 29-05-2001 20-08-2002 27-02-2001
WO 0023458	A 27-04-2000	AU EP WO	1318400 A 1123305 A1 0023458 A1	08-05-2000 16-08-2001 27-04-2000
US 5723598	A 03-03-1998	US US AT AU CA DE DE DK EP ES JP WO	5573905 A 6060596 A 193561 T 685050 B2 3944993 A 2132103 A1 69328781 D1 69328781 T2 643778 T3 0643778 A1 2147197 T3 7505530 T 9320242 A1	12-11-1996 09-05-2000 15-06-2000 15-01-1998 08-11-1993 14-10-1993 06-07-2000 26-04-2001 07-08-2000 22-03-1995 01-09-2000 22-06-1995 14-10-1993